NUMBER IN THE WORLD'S LANGUAGES

A COMPARATIVE HANDBOOK

Edited by Paolo Acquaviva and Michael Daniel

COMPARATIVE HANDBOOKS OF LINGUISTICS

Paolo Acquaviva and Michael Daniel (Eds.)
Number in the World's Languages

Comparative Handbooks of Linguistics

Edited by
Andrej Malchukov and Edith Moravcsik

Volume 5

Number in the World's Languages

A Comparative Handbook

Edited by
Paolo Acquaviva and Michael Daniel

DE GRUYTER
MOUTON

Contents

VI **Further perspectives on linguistic diversity**

List of authors

Paolo Acquaviva, University College Dublin, E-mail: paolo.acquaviva@ucd.ie

Wayan Arka, The Australian National University/Universitas Udayana,
E-mail: wayan.arka@anu.edu.au

Xavier Bach, Trinity College, University of Oxford, E-mail: xavier.bach@trinity.ox.ac.uk

Irina Bagirokova, HSE University and Institute of Linguistics, Russian Academy of Sciences,
E-mail: ibagirokova@yandex.ru

Svetlana Burkova, Novosibirsk State Technical University, E-mail: burkova_s@mail.ru

Denis Creissels, University of Lyon, E-mail: denis.creissels@univ-lyon2.fr

Mary Dalrymple, University of Oxford, E-mail: mary.dalrymple@ling-phil.ox.ac.uk

Michael Daniel, HSE University, Russian Federation, E-mail: misha.daniel@gmail.com

Gerrit J. Dimmendaal, University of Cologne, E-mail: gerrit.dimmendaal@uni-koeln.de

Abdelkader Fassi Fehri, Mohammed Bin Zayed University for Humanities,
E-mail: abdelkader.fehri@mbzuh.ac.ae

Elizaveta Filimonova, Rambler&Co, E-mail: ev.filimonova@list.ru

William A. Foley, University of Sydney/Columbia University,
E-mail: williamaugustefoley@gmail.com

David Gil, Max Planck Institute for Evolutionary Anthropology, E-mail: gil@shh.mpg.de

Ekaterina Gruzdeva, University of Helsinki, Institute of Linguistics, Russian Academy of Sciences,
E-mail: ekaterina.gruzdeva@helsinki.fi

Valentin Gusev, Universität Hamburg and Institute of Linguistics, Russian Academy of Sciences,
E-mail: vgoussev@yandex.ru

Daniel Harbour, Queen Mary University of London, E-mail: d.harbour@qmul.ac.uk

Angelika Jakobi, University of Cologne, E-mail: jakobi_a@web.de

Vadim Kimmelman, University of Bergen, E-mail: vadim.kimmelman@uib.no

Olga Krasnoukhova, Leiden University, University of Antwerp,
E-mail: o.krasnoukhova@hum.leidenuniv.nl

Yury Lander, HSE University, Russian Federation, E-mail: yulander@yandex.ru

William B. McGregor, Aarhus University, E-mail: linwmg@cc.au.dk

Andrew McKenzie, The University of Kansas, E-mail: andrew.mckenzie@ku.edu

Marianne Mithun, University of California, Santa Barbara, E-mail: mithun@linguistics.ucsb.edu

Paul Phelan, HSE University, E-mail: pphelan@alumni.nd.edu

Michinori Shimoji, Kyushu University, E-mail: michinoris@gmail.com

Adrian Stegovec, University of Connecticut, E-mail: adrian.stegovec@uconn.edu

Edward Vajda, Western Washington University,
E-mail: edward.vajda@wwu.edu

Viveka Velupillai, Justus-Liebig-Universität Gießen,
E-mail: viveka.velupillai@anglistik.uni-giessen.de

Valentin Vydrin, INALCO – LLACAN (CNRS) – St. Petersburg State University,
E-mail: vydrine@gmail.com

Beáta Wagner-Nagy, Universität Hamburg, E-mail: beata.wagner-nagy@uni-hamburg.de

Roberto Zariquiey, Pontificia Universidad Católica del Perú, E-mail: rzariquiey@pucp.edu.pe

Paolo Acquaviva and Michael Daniel

1 Number in the World's Languages: Configuring the variation space

1 Aims and context

How do languages express the category of *number*? What phenomena fall under this category, and how is this empirical domain to be circumscribed and structured? How is it related to other linguistic means of expressing quantity, such as quantifiers and cardinal numerals? What conceptual tools are best suited to describe this domain and to predict the systematic patterns that emerge, including what apparently cannot be expressed in a natural language? What generalizations emerge? These distinct but related questions, as well as others (like the question as to what number can ultimately 'mean' as a linguistic category), are certainly not new. On the contrary, together with the reflections on number that developed from the philological study of classical and Indo-European languages (e.g. Delbrück 1883), theorists like Humboldt (1830), Jespersen (1924) and Lyons (1968) recognized the subtlety and the typological diversity of the expression of number in natural language, focusing on some of the parameters of its variation. Number was the focus of Paul Forchheimer's pioneering study of pronominal systems (1953), one of the first sample-based typological studies ever; and it was one of the categories prominently covered by Greenberg's (1963) statement of linguistic universals (including 34, 35, and 36). Moreover, the complex network of phenomena linked to the expression of number (distributive and collective readings, countability, reciprocity, genericity, pluractionality) has been at the forefront of formal semantic research at least since Link's (1983) analysis of plurality in connection with countability.

And yet, Greville Corbett could with reason begin his landmark study (Corbett 2000) noting that number is "the most underestimated of the grammatical categories", despite many insightful observations scattered in the literature (the references to Jespersen and Lyons above are taken from Corbett's opening lines). The main reason was the lack of comprehensive accounts that could do justice to the impressive, often unsuspected and, as the chapters that follow suggest, still underestimated diversity of number systems across languages, analyzing number from a unified perspective. Corbett's (2000) monograph contributed to precisely this aim, and it is no exaggeration to say that in doing so Corbett marked a milestone in the history of research on number in language. His work has informed all subsequent research across frameworks (a fact prominently recognized by Harbour 2014, in a major contribution to the formal semantic study of the category), and it provided the impulse for a broad range of typological overviews such as the number-related chapters in the World Atlas of Language Structures (Dryer and Haspelmath 2013, specifically chapters 33, 34, 35, 36). Since Corbett's study of number, not only has this category

enjoyed a renewed interest from the typological perspective (Moravcsik 2017, Aikhenvald 2018, Ivani & Zakharko 2019), but also from more formally oriented work, especially – though not exclusively – in connection with the analysis of countability (Chierchia 1998; Sauerland 2003; Borer 2005; Harbour 2007, 2014; Nevins 2011; Acquaviva 2008; Alexiadou 2011; Massam 2012). The advances made in the last two decades in these various directions justify new syntheses, such as the recent publication of Cabredo Hofherr and Doetjes (2021), who examine the conceptual common ground shared by various approaches, reassess traditional generalizations, reappraise the empirical landscape including lesser-known phenomena, and broaden our view of typological diversity.

Our goal is not to achieve all of this in one book, which would be an unrealistic ambition, but to offer a substantive empirical contribution towards such a broad-ranging new synthesis. The chapters that follow are devoted to one language each, sometimes against the background of its sister languages, or to a small family of closely related languages; the two exceptions – the one which considers South American languages, and the other which considers contact languages – still circumscribe a set of systems where the expression of number is by and large mutually comparable. This reflects the choice that lies at the basis of the other volumes in this series: the grammatical topic under investigation (here, number) is presented and analyzed in a wide range of the empirical aspects it presents across linguistic systems; and it organizes the respective notional, functional or cognitive domain, depending on the framework.

This domain seems to lack a dedicated and cross-framework label that would be seen as conventionally paired with the linguistic category of *number*, such as when *mood* is paired with modality (possible or necessary states of affairs, volition and other), *tense* and *aspect* with temporality (time reference) and topological relations such as anteriority or inclusion, or *person* with speech act participants. The concept of quantity, which is sometimes used in this sense, is broader than the notional domain subsumed under the linguistic category of number, since language can quantify amounts that cannot be represented on a numerical scale (as in *there was a lot of bitterness in his words*, *a little effort is required*), and because the notional domain of quantity includes comparison and relative size, which are not, to the best of our present knowledge, expressed by linguistic number and do not have to involve magnitudes that are numerically specifiable (*this is more journalism than scholarship*, *I feel sadder now than an hour ago*). On the other hand, notions such as abundance or scarcity can find linguistic expression through the number category (think of plurals of abundance or greater plurals and paucals in Corbett 2000), but these are marginal values of the category, attested in some systems but comparatively rare, both in terms of corpus occurrences in individual languages and in terms of cross-linguistic types. They seem to be subordinate to the main opposition, centred on the contrast between being one and being more than one (what can be called unity vs. multiplicity). Nor is number reducible to the notion of cardinality

because the information encapsulated in number values references cardinality insofar as it distinguishes one (which can be seen as unity rather than cardinality) or, less frequently, two (cardinality of two), or very rarely three (cardinality of three) from more, but never distinguishes between higher cardinalities. This function is accomplished by numerals, another linguistic means to express quantity which interacts with number but is not number.

Number appears to be a primitive notional dimension, rooted in the foundational notion of identity as evidenced by the philosopher Peter Simon's pithy statement that "where there is identity there is difference, and where there is difference there is number" (Simons 1987: 151). In other words, number as a cognitive dimension rests on the key distinction between what is one and what is not one and is directly rooted in the fundamental opposition between being one and the same and not being one and the same. Unity and multiplicity, as well as size distinctions based on different values of multiplicities, arise conceptually from the foundational notion of identity; these notions are intimately interrelated conceptually, and they define without intermediaries a cognitive dimension that we simply call 'number'. Correspondingly, it is normal to talk of the 'number sense' (Dehaene 1987) which underlies both the ability to estimate the magnitude of a group at a glance, and the mathematical skills that rely on a symbolic representation. We can then broadly understand 'number' as the range of notional characterizations related to unity and multiplicity, potentially extending to some cardinalities along a numerical scale (duals, trials and greater plurals), or reducing such a numerical scale to a qualitative scale of magnitude (paucals). We can then define the goal of this volume as a contribution to better understanding the ways in which human language represents and organizes information about number in different grammatical systems.

Each chapter of the volume explains how number works in a grammatical system as a whole. The way this account is structured is the same for all chapters, as it is based on a questionnaire, so as to simplify typological comparison of parameters across chapters. At the same time, however, this is not a typological overview of number systems from one particular perspective translated into a set of questions. The chapters can be read as self-contained case studies because they do not limit themselves to recording how a given system behaves with respect to a pre-established set of parameters. While we came up with a selection of phenomena that we think may be cross-linguistically relevant and delivered it to the authors in the form of a list of topics to cover, our intention was that the authors should take the questionnaire as a point of departure, rather than be caught in a Procrustean bed of parameters. The contributors were invited to offer their interpretation of the number system of a language, emphasizing the coherence internal to the system. When compiling the questionnaire, we did not directly specify one or a set of big theoretical questions but rather aimed at making the accounts detailed enough to naturally address these questions and if necessary challenge the expectations of the readers (and indeed of us editors). The overall goal was not so much a full typology

of number systems, as an extensive phenomenology of number, where each phenomenon is placed in the full context of its linguistic system and the single accounts are coordinated with each other to facilitate direct comparability. In short, the aim is a presentation of number in a reasonably diverse sample of the world's languages.

2 The sample

Inevitably for a book-sized collection of detailed accounts of a linguistic category in individual languages, we cannot hope for our sample to be representative of the typological spectrum in a way that comes anywhere close to exhaustivity. We could not sample our languages on purely structural grounds because there is no known connection between the typological profile of a language as a whole and the variation of number systems. Most probably, there is no such connection at all; at least, this must be the null hypothesis. One cannot say whether, for instance, agglutinative languages are more or less inclined to possess the category of dual; and most probably they are not. The only typological bias we had was towards languages with some expression of number outside the pronominal domain, which correlates with a 'higher than average' degree of morphological complexity. But on the scale of morphological complexity, our sample covers very different cases, from polysynthetic languages like Circassian (Northwest Caucasian) to languages with relatively poor morphologies like Eastern Dan (Mande) or Indonesian, where number, when expressed at all, is expressed by plural words or reduplication.

Another relevant factor has been the availability of specialists who could provide grammatical accounts based on direct engagement with the data and the literature. The chapters present a wealth of data, often original and first-hand, selected by language experts who are best qualified to interpret the phenomena in the context of the grammar of the respective languages.

Within these limits, we made a reasonable attempt to achieve diversity in terms of geographic and genealogical balance. Again, given practical constraints on the number of chapters, we could not possibly use a representative sample but only a relatively balanced one. This is visualized on the map (courtesy George Moroz), with black dots corresponding to chapters on individual languages (or small groups of related languages) and the white dots in South America representing the 100+ languages sample used in Krasnoukhova's chapter (this volume).

As the map in fig. 1 shows, we had two to three languages from what may be considered large geographic areas, including Tswana (Bantu), Eastern Dan (Mande) and Karko (Nilo-Saharan) from Africa; Occitan (Romance) and Slovenian (Slavic) from Europe; Arabic (Semitic) and Circassian (Northwest Caucasian) from the Middle East and the Caucasus; Nganasan (Uralic), Ket (Yeniseian) and Nivkh (isolate) from northern Asia; Japonic languages and Indonesian (Astronesian) from the Pacific; Marori (Trans-New Guinea), Sepik languages and Gooniyandi (Bunuban) from
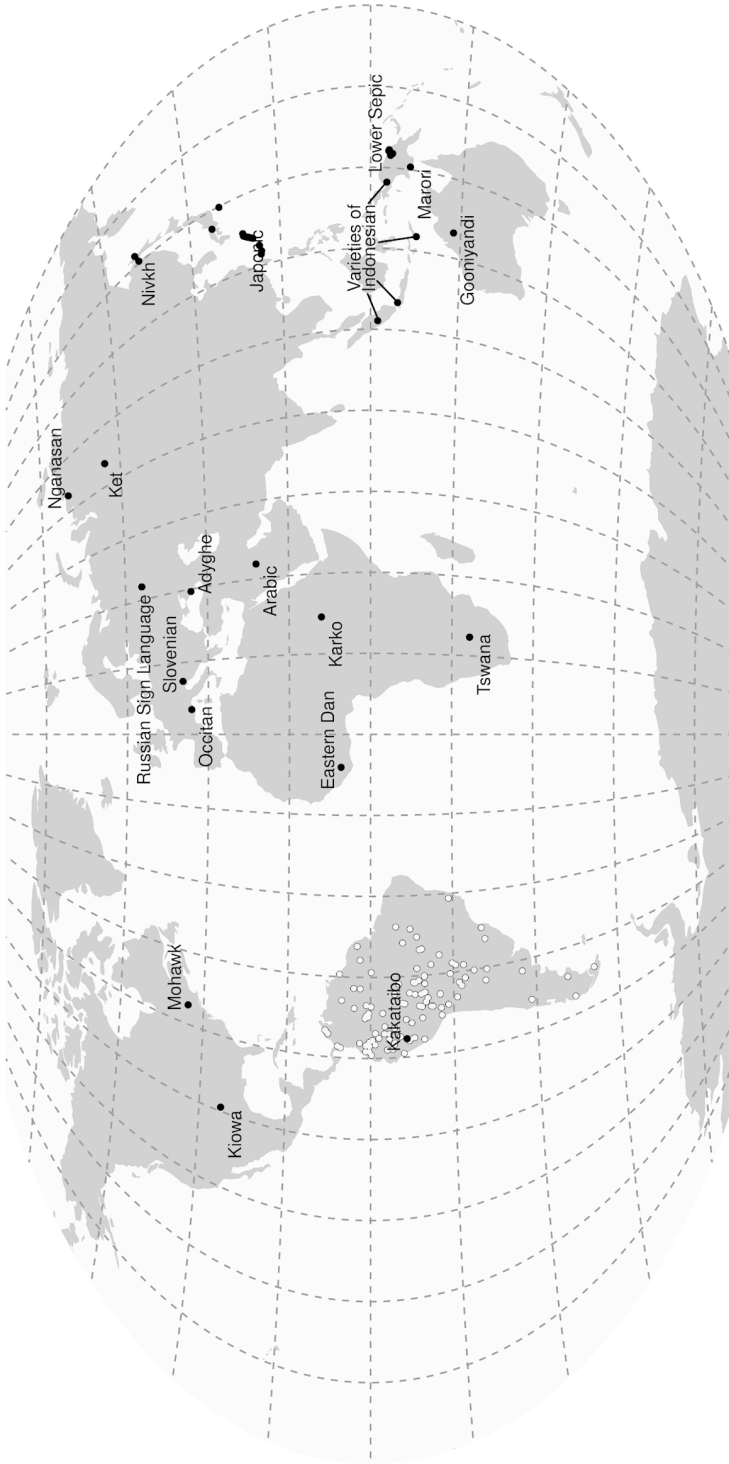
Fig. 1: Language sample (Map created by George Moroz using the lingtypology package [Moroz 2017] for R.).

New Guinea and Australia; Kiowa (Kiowa-Tanoan) and Mohawk (Iroquoian) from North America; and Kakataibo (Panoan) from South America. Within each area, only Occitan and Slovenian in Europe, and Tswana and Eastern Dan in Africa, are genealogically related, and even these only distantly (Indo-European and the contested Niger-Congo, respectively). All other languages are not related, although some chapters include a background comparison with the languages of the same family, including the chapters on Karko and Ket; and some offer surveys of families rather than individual languages, including Lower Sepik and Japonic. One chapter is an areal survey of the languages of South America and one is an overview of contact languages. Finally, one chapter is an analysis of number in Russian Sign Language, also offering some background on other sign languages.

Because of the limited number of languages covered, the book cannot provide any estimates of the comparative frequency of the pre-established values of different parameters of number systems, addressing questions such as 'How common is the dual?' or 'How often does the use of a numeral rule out morphological expression of number?', the kind of questions that the World Atlas of Language Structures addresses (Dryer & Haspelmath 2013). But giving our contributors relatively free rein in their treatment of the questionnaire led to unearthing novel details, some of which are discussed in the concluding chapter of this volume. Typologically, these insights may at present remain anecdotal, possibly reflecting an idiosyncratic feature of an individual language system. But they open new perspectives for those who work on language description. In short, our aim was not so much to provide new generalizations about old parameters but to look for new parameters that may call for novel generalizations.

In sum, the chapters that make up the bulk of this book present a great variety of phenomena concerning the grammar of number, taken from a genealogically and geographically diverse sample of grammatical systems. Each chapter illustrates the workings of number in a language as a whole, highlighting phenomena of particular interest but always presenting the full picture. Taken together, they constitute an empirically oriented encyclopaedia of number systems, or at least a fragment thereof. But the individual chapters would not constitute an organic whole if they were not conceived and structured according to the same unifying plan, which is expressed in the questionnaire that guided the writing of each contribution.

3 The questionnaire

This volume was conceived as a comparative handbook. Every comparison requires a frame of reference, and any frame of reference inevitably shapes the results of the comparison. Accordingly, our questionnaire defines a certain structure and asks certain questions, and in doing so it partly shapes the answers it receives. More strongly, it could be seen as laying down a hypothesis about the possible variation

space of number systems in natural language. While being aware of this, but without deliberately striving to formulate such a hypothesis, we have sought a compromise between the goals of maximum empirical coverage (by giving our authors relatively free rein as to what they may want to include in the discussion) and ensuring comparability across chapters (by providing a relatively rigid general structure and a typologically informed wish-list of questions as guidelines to all chapters). We also promoted the use of a descriptive metalanguage that is as inclusive and as theoretically neutral as possible.

The fact that each chapter responds to the same questionnaire and is structured accordingly is in line with other volumes in the same series. The questionnaire, here reproduced on p. 15–24, is far from being a checklist. Instead of verifying the presence or absence of certain phenomena, or asking which option is attested among a restricted set of choices, or requiring a dataset of stimuli translated into all languages, it has first and foremost the task of structuring and streamlining the authors' accounts of the number system by providing a conceptual skeleton for the chapters. Every chapter begins with an *overview* section giving historical and sociolinguistic background on the language, the language's typological profile, and generalities about its number system. It is followed by a section on the *formal expression* of number distinctions in pronouns, nouns, verbs, and eventually other word classes. Next is a section on the ways number is relevant to (*morpho*)*syntax*, most prominently agreement in number, and then a section on the interplay between number and *semantics*, *pragmatics* and *discourse*, which may host an assorted range of issues from politeness in pronouns to referentiality.

The general overview places the language in its typological context, outlines the key properties that characterize its number system as a whole, and helps the reader to find one's path through the chapter. This global viewpoint is taken up again in the conclusion, which comments on the system just described and points to aspects of particular interest or in need of further analysis. Broadly descriptive rubrics like 'pronominal number', 'agreement and syntax' or 'semantics and discourse' provide a uniform but rough division of the subject matter. Within them, each chapter does not list the number-related phenomena according to a pre-established list, but presents them following the author's interpretation of the way number is expressed in the relevant domain of grammar and as part of the linguistic system being analyzed. Yet, even this coarse articulation might assume distinctions that have no basis in the empirical reality of a language. For example, Gil's account of pronominal number in Indonesian in this volume starts by making it clear that the distinction between nouns and pronouns is based on semantic properties (pronoun forms rigidly express the person values for which they are specified) but cannot be identified in morphology or syntax, as lexical forms taken by themselves underdetermine the word-class status they take on in any given construction. In addition, for three of the four varieties of Indonesian described by Gil, the class of pronouns is 'semi-open' (the list of its members is in constant flux), and while in

certain contexts number must obligatorily be expressed by using the corresponding form where a paradigmatic opposition is available, varieties differ noticeably in when this choice is mandatory, for which persons it is available, and what values are specified. Moreover, to put it in a way directly relevant to number itself, the difference between nouns and pronouns may be based solely on the treatment of number distinctions, as in Creissels' account of pronominal number in Tswana (this volume). Creissels sums up the main difference within the nominal system with the terse statement that "the only asymmetry between nouns and pronouns concerns plural exponence in first and second person pronouns". In yet other cases, while the difference between nouns and pronouns is not questioned, it may seem (almost) irrelevant in terms of the expression of number, as in Kakataibo (Zariquiey, this volume). Finally, in terms of their relation to number, some 'nouns' may be intermediate between pronouns and regular nouns, as 'address nouns' in Japonic languages (Shimoji, this volume), suggesting that the coarse distinction between nouns and pronouns may veil finer distinctions along the Animacy Hierarchy. In short, in individual linguistic systems, the boundary between pronouns and nouns, as solid as it is given in the questionnaire, may either be less prominent or more complex than suggested.

As another example of problematic mapping between the questionnaire and individual languages, some authors, especially those working on polysynthetic languages, strongly – and justly – objected to calling 'agreement' the pronominal marking on the verbal complex. In languages like West Circassian (Bagirokova et al., this volume), such elements are more adequately described as bound pronouns expressing verbal arguments, rather than agreement understood as morphosyntactic covariation with an external NP. Under this view, what ends up in the section on syntax as description of agreement could have been equally well (or more appropriately) described, for this and similar languages, in the section on the expression of number (as pronominal number). More relevantly to the typology of number, the same authors also discuss cases of lack of isomorphism between the expression of number on arguments and predicates, arguing that the number morphology on the predicates may be a device (partly?) independent of the expression of number on their arguments. Similarly, McGregor argues that the expression of argument number on Gooniyandi verbs is not agreement with the nominal arguments but a distinct linguistic encapsulation of number inherent in verbs. The phenomenon is not agreement; but we need the notion of agreement to see how it departs from it.

Similarly, notions like 'mass / count', 'classifiers', 'associative plural', which feature in the finer-grained questions that flesh out the broad categorization of the questionnaire, and indeed the very categories of nouns, pronouns, and verbs (and the related notions of nominal, pronominal, and verbal number), have a purely practical value in the questionnaire. We may consider them 'comparative concepts', not 'descriptive categories' (or etic rather than emic categories, as extensively discussed by Gil in application to associative plurals and inclusory constructions).

They structure our thinking about the variation space, and as such are not (nor do they purport to be) 'neutral' or 'directly justified by the empirical evidence'; but the responses that they elicit may well relativize or problematize their content. Indeed, the very notions of 'singular' and 'plural' may necessitate a theoretical contextualization, as in Fassi Fehri's distinction between singular-plural and collective-singulative oppositions in his treatment of Arabic (this volume).

All such instances of tension between the categories named in the questionnaire and the phenomena described in the chapters were resolved by keeping the larger section titles and structure for the sake of cross-chapter comparability, but leaving it to the full discretion of the authors to refute the misnomer straight in the opening of the section, suggesting more neutral or more adequate terms (such as indexation instead of agreement in the case of West Circassian), or explicitly invalidating the division imposed by the questionnaire (such as the conventional part-of-speech classification in the case of Indonesian, as we have seen). In some cases, the authors could simply indicate that the parameter has been treated in a different section; is inapplicable to the language; or has not been sufficiently studied.

We have illustrated the potential empirical incongruity of the questionnaire with only two examples of assumptions built into it: the division between nouns and pronouns and the understanding of expression of argument number on the verb as syntactic agreement. More such examples are dispersed across the chapters. While suggesting to the authors a rigid structure to follow may seem like sacrificing empirical soundness for the sake of ease of use of chapters for the purposes of typological comparison, at the end of the day, we believe that the necessity to structure the description along a general questionnaire, while not equally welcomed by all authors, may have been ultimately beneficial to the volume also in a deeper sense. Indeed, the effect of this approach is that it foregrounds the deviations of individual linguistic systems from our typological (and we must admit, in part Eurocentric) expectations. This is similar to how our current understanding of the notions of both clusivity and first person plural is based on a careful reassessment of the category of clusivity as viewed through the prism of European languages used as descriptive metalanguages (Filimonova 2005).

We thus contend that, despite the structural uniformity imposed by the questionnaire, each chapter is a coherent analysis of the number system in a language, according to the logic of that system. The typological profile of the systems in question is first made explicit in the opening of each chapter. The data are presented against this background; the 'isolating', 'monocategorial', and 'associational' profile of Indonesian (Gil, this volume) can thus be contrasted with the classic noun-class profile of Tswana (Creissels, this volume). Within each typological profile, the logic behind the organization of number phenomena in a given system is always a matter of interpretation, and, to the extent that this interpretation deviates from expectations incorporated into the questionnaire, the authors may take a clear stand regarding their analytic and theoretical choices. Every chapter is 'descriptive' in the

sense that its aim is to present the most accurate interpretation of a given system rather than to establish some theoretical results. They are 'framework-neutral' in a way we believe was almost maximally achievable in a co-operative project with many contributors, and certainly fully accessible to readers regardless of their theoretical persuasions; but they are all *analyses* in their own right.

What is probably most striking is the extreme diversity emerging from this sample that is in no way diminished by the unifying frame imposed by the questionnaire nor, as it seems, by the lack of principled sampling. Not a single system displays an overall similarity to any of the others, and each shows patterns that are not reiterated in the rest of the sample – or, if they are, they are instantiated in a very different way, in a language with a very different structure, with common functional patterns masked by apparent structural differences.

On the similarity side, consider reduplication. One of its oft-quoted functions is the expression of plurality (Corbett 2000: 148; Rubino 2013), and it has been observed, including in the works just cited, that reduplication may be associated more specifically with distributivity. And yet it is impressive how this specific association reappears again and again in our sample, across continents, parts of speech and even modalities, with reduplication and distributivity working in tandem to become a prominent if not the dominant type of plurality in sign languages (Pfau & Steinbach 2006, also Kimmelman et al., this volume), thus probably corroborating the view that distributivity is not simply a subtype of plurality but an independent category (Corbett 2000: 111–120).

Another, less a priori obvious case in point is provided by comparing Kiowa (Kiowan-Tanoan) and Karko (Nilo-Saharan). Kiowa features the crosslinguistically rare pattern of inverse number (in fact, Corbett's overview of inverse number (2000: 159–166) opens with a discussion of Kiowa), where the same marker is used for the singular or the plural, depending on lexical semantics of the noun. Karko shows a system typical of this language family, with nouns differing as to which number values will be marked (plural, singular, both, or none). While the systems and the terms used to refer to them are different, the comparison of the underlying nominal classification shows properties common to the two languages.

As to the typological diversity of number, it is naturally easiest to illustrate it by looking at the variation in the types of its exponence, and our sample happens to be illustrative in this respect. In Tswana, in the south of Africa, the expression of nominal number is cumulated with gender in a system of multiple prefixes typical of Bantu. It also does so in Lower Sepik languages of northern Papua New Guinea, where the two categories are cumulated in suffixes; and in Marori, an Austronesian language of southeastern Papua New Guinea, these are coexpressed in harmonic vowel alternations. While in these languages number markers abide through complex patterns of cumulation with gender, Yenissean derives all its intricate allomorphy of plural exponence from only two morphemes, *-n* and *-ŋ*, whose synchronic distribution is partly allomorphic and partly motivated by the Animacy Hierarchy.

Occitan, a Romance variety (or rather a family of varieties) which might have been expected to yield few surprises, also displays an impressively diverse range of exponents, which include alternation in stress and in prosody (low pitch on plural endings was attested in one particular variety). To express number, Gooniyandi in northern Australia and Kakataibo in the Amazons use clitics, and Eastern Dan in West Africa uses a prosodically independent plural word. While analytic expression of number is also commonly thought to be typical of contact languages, this is not uncontroversial, as discussed by Velupillai (this volume). Finally, in Modern Arabic, the multiple strategies of pluralization range from cross-linguistically common affixation to the transfixation almost unique to Semitic, which adds to the typological palette of formal complexity of number exponence.

But there are many other ways on top of exponence, from more to less expected, in which languages may differ in their treatment of number.

The category of nominal dual is attested in various languages of our sample, including in Nganasan and in Ket, in the north of Eurasia, in Slovenian, in central Europe, in Arabic, in northern Africa and the Middle East, in Kiowa, in North America, and in Gooniyandi, in northern Australia. But in Ket, the (rather marginal) duals are of the 'ambal' kind (Plank 1989), while Nganasan, Arabic and Slovenian can or tend to refer to naturally paired referents like body parts by using plural forms. In Gooniyandi, the dual clitic signals a 'cohesive' (collective) interpretation, and so does the plural clitic; in Slovenian, the plural but not the dual encodes the collective reading; and in Kakataibo collectivity is expressed by special suffixes that may combine with plural clitics.

The expression of verbal number by stem alternation across areas and genealogical units described by Krasnoukhova (this volume) for South America makes this chapter perfect grounds for a comparison with less dense instantiations of this category in the languages of other continents, as in Marori in Papua New Guinea (dozens of verbs) and Ket in Siberia (quite a few verbs). But even more importantly, against this background, it seems unexpected and calling for an explanation that in Eastern Dan, of all lexical items that are semantically predicative, expression of plurality by stem alternation is limited to several adjectives.

Indonesian, Nivkh in eastern Siberia and various languages of South America all share numeral classifiers, but while Indonesian by and large complies to the stereotypical image of a classifier language in that it essentially lacks morphological nominal plural, Nivkh features relatively frequent morphological plurals, and the many classifier languages of South America combine numeral classifiers with obligatory morphological pluralization in the same numeral phrase (see also Seifart and Payne 2007).

In terms of lexical distribution of number, both Japonic languages and varieties of Indonesian feature a special class of terms of address, and in both cases these nouns show idiosyncratic behaviour with respect to number. But while Japonic terms of address show obligatory number marking, which aligns them with personal pronouns, Indonesian terms of address in this very function cannot be pluralized.

And to give one final example of diversity masked by formal similarities across number systems, also in lexical terms but in a more phenomenological vein, Indonesian of the Pacific, West Circassian in the Caucasus and Nivkh in Siberia, and Mohawk of the northeastern US problematize the notion of nominal vs. verbal number by using the same marking on both. But in Mohawk, in accordance with what has been suggested as a frequent path of evolution of number in North America in Mithun (1988), the number marking (for humans) is smuggled into nominal morphology together with deverbal nominalizations; while in Nivkh the nominal plural marking on the predicate has probably appeared through the generalization of the use of nominalizations as finite predicates; and Indonesian with its flexible part-of-speech system simply applies its only morphological pluralization device, reduplication, indiscriminately to both arguments and predicates.

To sum up, the questionnaire is designed as a list of questions organized into general sections that remain the same for all chapters so as to allow a comparison of each aspect of the number system across languages, but always with the phenomena firmly placed in the context of the respective language.

Acknowledgements

References

Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford University Press.

Aikhenvald, Alexandra. 2018. *Number systems in grammar*. Initial Orientation paper for an LCRC Global Workshop 'Number systems in grammar', Cairns: LCRC.

Alexiadou, Artemis. 2011. Plural Mass Nouns and the Morpho-syntax of Number. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th West Coast Conference on Formal Linguistics*, 33–41. Somerville, MA: Cascadilla Proceedings Project.

Borer, Hagit. 2005. *In Name Only*. Oxford: Oxford University Press.

Chierchia, Gennaro. 1998. Plurality of Mass Nouns and the Notion of Semantic Parameter'. In Susan Rothstein (ed.), *Events and Grammar,* 53–103. Dordrecht: Kluwer.

Corbett, Greville. 2000. *Number*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Dehaene, Stanislas. 1987. *The Number Sense*. Oxford: Oxford University Press.

Delbrück, Berthold. 1893. *Grundriss der Vergleichenden Grammatik der Indogermanischen Sprachen*. 3nd volume. Syntax der Indogermanischen Sprachen, Part 1. Strassburg: Tübner.

Dryer, Matthew & Martin Haspelmath. 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Filimonova, Elena (ed.). 2005. *Clusivity: Typology and case studies of inclusive-exclusive distinction*. (Typological Studies in Language 63). Amsterdam: John Benjamins.

Forchheimer, Paul. 1953. *The Category of Person in Language*. Berlin: Walter de Gruyter.

Greenberg, Joseph H. 1963. *Universals of language*. Cambridge and London: The MIT Press.

Harbour, Daniel. 2007. *Morphosemantic Number: From Kiowa Noun Classes to UG number features*. Dordrecht: Springer.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90. 185–229.

Humboldt, Wilhelm von. 1828. Ueber den Dualis. Druckerei der Königlichen Akademie der Wissenschaften.

Ivani, Jessica & Taras Zakharko. 2019. *Tymber, Database on nominal number marking constructions*. Version 1.0.0. https://github.com/jivani/tymber/

Jespersen, Otto. 1992. The philosophy of grammar. University of Chicago Press.

Link, Godehard. 1983. The logical analysis of plurals and mass nouns: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language,* 303–323. Berlin: De Gruyter.

Massam, Diane. (ed.). 2012. *Count and Mass across Languages*. Oxford: Oxford University Press.

Mithun, Marianne. Lexical categories and the evolution of number marking. Theoretical morphology: Approaches in modern linguistics 21 (1988). 1-234.

Moravcsik, Edith. 2017. Number. In Alexandra Y. Aikhenvald & R.M.W Dixon (eds.), *The Cambridge handbook of linguistic typology,* 440–476. Cambridge: Cambridge University Press.

Moroz, George. 2017. Lingtypology: Easy Mapping for Linguistic Typology. <URL: https://CRAN.R-project.org/package=lingtypology>

Nevins, Andrew. 2011. Marked targets vs. maked triggers and the Impoverishment of the dual. *Linguistic Inquiry* 42. 413–444.

Pfau, Roland & Markus Steinbach. 2006. Pluralization in sign and in speech: A cross-modal typological study. *Linguistic Typology* 10(2). https://doi.org/10.1515/LINGTY.2006.006.

Plank, Frans. 1989. On Humboldt on the dual. In Roberta Corrigan, Fred Eckman & Michael Noonan (eds.), Linguistic categorization. *Proceedings of an International Symposium in Milwaukee, Wisconsin, April 10–11, 1987*, 293–333. Amsterdam: John Benjamins.

Rubino, Carl. 2013. Reduplication. In Matthew S. Dryer & Martin Haspelmath (eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/27, Accessed on 2021-03-31.)

Sauerland, Ulrich. 2003. A new semantics for number. In Robert Young & Yuping Zhou (eds.), *Proceedings of Semantics and Linguistic Theory SALT 13*, 258–275. Cornell, Ithaca: CLC Publications.

Seifart, Frank & Doris Payne. 2007. Nominal classification in the North West Amazon: Issues in areal diffusion and typological characteristics. *International Journal of American Linguistics* 73. 281–287.

Simons, Peter. 1987. *Parts: A Study in Ontology*. Oxford: Clarendon Press.

1.A Number in the World's Languages: Questionnaire

The following questionnaire assists authors in structuring their chapter in a way that ensures that all chapters consider a broad range of phenomena and are easily comparable.

The introduction serves to provide a general background on the language (language family), including its geographical and sociolinguistic situation and typological profile, and to briefly outline the position of the language in the typological spectrum relative to number. The conclusion is open-ended and interprets the global picture described in the preceding sections, including open questions and contentious issues.

Not all the topics will be relevant for all chapters; but if a phenomenon is certainly or very likely absent from the language, it is desirable to state it explicitly. This does not mean, however, that all single bullet points should be mentioned explicitly; they only guide the composition of the chapter. The unity of structure is achieved by following the numbered sections and subsections.

The parts indented by a dash complement the topic by expanding on what kind of empirical evidence could be relevant to the discussion in the (sub)section; these are guidelines and suggestions for answering the general questions in the most informative way.

If you envisage using this questionnaire in your research, we recommend that you consider the general discussion of epistemological issues with its application to individual languages in the introduction and the conclusion to this volume, as well as the ways it was eventually applied by the contributors to their data. Note that some specific phenomena and parameters of cross-linguistic variation came into focus during our collective work on the volume. To make this questionnaire useful for future data collection and fieldwork, we have included these into this printed version. Note, however, that not all authors addressed them in their chapters.

Outline of the questionnaire

1 Overview of the number category in the relevant language / language family
2 Pronominal, nominal, and verbal number
   2.1 Generalities
   2.2 Pronominal number
   2.3 Nominal number
   2.4 Verbal number
3 Agreement and the syntax of number
4 Semantics and discourse
5 Conclusions

1 Overview of the number category in the relevant language / language family

The introductory section contains background information on the language or language group and briefly delineates its typological profile with respect to number.

- geographic, genealogical and sociolinguistic background
- typological profile
    - cumulative vs. agglutinative vs. isolating?
    - are there classifiers/noun classes (genders)?
    - is the presence of a number contrast typologically unexpected or notable? (e.g. in a classifier language)?
- status of number generally
    - obligatoriness, inflectional or non-inflectional status; affixation or plural words; regularity and productivity of exponence.
- number values
    - are they always the same across lexical categories (parts of speech)?
    - are there approximative values like *paucal*?
    - are there prominent *collectives*?
    - is there a singulative-collective opposition? *plurative* as a secondary pluralization of *singulative*?
    - are there *associative plurals* ('X and the X-related group') and *dyadic kinship terms* ('people in the specified kin relation to each other)?
- major differences between categories, especially nouns and pronouns (reference to 2.2–2.3)
- number splits across the lexicon
    - is there any kind of major lexical / semantic splits within nouns (reference to 2.3)?
    - are there widespread doublets (semantically differentiated alternative plural forms for the same lexical item) (reference to 2.3)?
- is there a (traditionally recognized) *verbal number* as distinct from agreement (reference to 2.4)?
- status of property words
    - if they are involved in number marking, please briefly characterize adjectives in terms of their part-of-speech status: are they morphosyntactically predicates or nouns rather than attributes?
- domains of agreement
    - is there any kind of clausal agreement? is there any kind of NP-internal agreement? if so, how (ir)relevant is number? (brief outline, with reference to section 3);
    - if relevant, accusative / ergative / other basis of agreement.

2 Pronominal, nominal, and verbal number

2.1 Generalities

–   major asymmetries between nouns and pronouns (e.g. range of values, semantics, exponence)?
–   evidence for *verbal number* as distinct from agreement?

Note: Further structure is partly the author's choice. Whether pronominal number and nominal number should be treated together, in the same section, is decided depending on whether there are major differences between them (see above). Whether special sections on e.g. interrogative pronouns are required, depends on how different these are from other pronouns / nouns, respectively.

2.2 Pronominal number

–   values and exponence on personal pronouns; presence of the inclusive value
     –   expression of pronominal number: suppletion? affixes? same as on nouns?
     –   if present, what is the status of dual inclusive (dual or Greenberg's minimal)?
     –   if relevant: are there any differences, in terms of values and exponence, between independent and clitic pronouns (and any other variants of personal pronouns)?
     –   how is number expressed in indexing personal arguments in the verb?
     –   expression of number on other classes of pronouns

Note: if necessary – from the point of view of exponence or otherwise – add further subsections on other subclasses of pronouns, including:
     –   interrogative pronouns;
     –   demonstrative pronouns, including both their adnominal and independent uses;
     –   other pronouns (logophoric / reflexive etc.).
     –   if applicable, describe optional number marking on pronouns and its factors
     –   if applicable, describe inclusory constructions ('Peter and I went' expressed literally as *we and*/*with Peter went* or (*with*) *Peter* [1Pl]*went*)

2.3 Nominal number

– typological profile
  – is number an inflectional category?
  – presence of gender / noun classes / (numeral) classifiers;
  – relevance of the Animacy Hierarchy (for exponence, for the set of values, for their optionality); any discrepancy from the Hierarchy?
  – values: any difference as compared to pronouns?
– exponence
  – interaction of number (on nouns) with other nominal grammatical categories: primarily, is there systematic cumulation (with gender, case, definiteness, or other categories)?
  – dominant patterns; distribution of different markers (not formally allomorphic, i.e. not automatically determined) across lexical bases;
  – plural words, i.e. free-standing morphemes that express numerosity (plural, dual, paucal etc., but not just 'many' or 'few') of a noun they accompany; their position in the NP / clause; if known at all, what is their prosodic status;
  – number suppletion (minimally – list of sample items, best – nearly exhaustive list);
  – extended exponence (more than one plural marking, e.g. Irish *ubh* 'egg', pl. *uibh-eacha* or *uibh-each-aí*, with the plural affix + plural end-vowel)
  – *constructed* aka *hybrid* number (e.g. singular on V and plural on N resulting in dual reading)?
– numeral classifiers
  – their inventory;
  – their use outside quantification contexts;
  – presence of *collective classifiers* (for groups, e.g. Tongan *ko e TAKANGA ika* 'a SHOAL of fish', Chinese *shi DUI luobo* 'ten PILES of carrots') as opposed to plural/non-unitizing classifiers (with non-singular reference, e.g. Mandarin *yī xiē shū* 'one CL book' = 'a few books', like *yī xiē shuǐ* 'some water');
  – eventual co-occurrence with number marking.
– interpretation of values
  – interpretation of forms unmarked for number: are these forms interpreted as strictly singular? can plural and eventually dual be considered as optional marking? under which conditions it occurs?
  – interpretation of plural: special readings of plural forms, or special marking used to this effect, including: abundance (e.g. 'lots of snow'), scattered (e.g. 'snow here and there'), sort plurals ('various types of N') – especially for mass nouns; associative plural ('X and his group');
  – any systematic asymmetry between *strict* (only two or more) and *non-strict* (possibly one) reading for plural nouns? (non-strict ex.: *are there oranges? yes* – as truthful answer if there is exactly one)

- are certain forms of plural incompatible with a generic reading, e.g. *bears hibernate*?
- interpretation of dual: can dual be used for natural dyads (pairs)? can it be used for anything else than natural dyads?
- any non-deterministic distribution (e.g. either paucal or plural can be used for a small collection; or dual and plural are in some way not in (a fully) complementary distribution).
- distribution across the lexicon
  - relevance of animacy; special lexical classes like measures, mass-nouns, units of time, abstract nouns, action nominals, place names and other names; semantics of non-singularity with such nouns
- special plural categories
  - any special plural forms for *collectives*? what exactly is, in your language, a collective reading of a plural?
  - special morphology for *associative plural* ('X and his group');
  - special morphology for *dyadic kinship terms* (special forms of kinship terms and construction involving them describing two or more people in this kinship relation between each other).
- any lexical classes / individual nouns that appear in one specific value with a *general* number reading?
  - e.g. noun forms that can be used indifferently for units or groups, like the English formally singular names for animals in *we hunted lion*, *we shot three lion*; or Miya (Chadic), where inanimate nouns may be interpretively plural but not so marked (Corbett 2000: 73).
- number deficiency in nouns: singularia, pluralia and dualia tantum
  - interaction between nominal morphology of such nouns (NP exponence) and agreement;
  - eventual semantic motivation of formal deficiency;
  - singularia tantum: what lexical domains resist pluralization; how strictly; what semantic effects are associated with their after-all-pluralization.
- mass and matter nouns
  - do (some) mass readings share markers with plural readings? can mass nouns be pluralized and remain mass? What are agreement patterns for mass nouns?
- group and set nouns
  - are there nouns with intrinsically plural reference? (cf. English *police*); nouns used for collections? (cf. English *flock*); metonymic uses of nouns like *village*? (as for village inhabitants); do they form morphological plurals? what is their agreement pattern (plural or singular for the morphological singular)? any other noteworthy properties related to number?
- action nominals
  - do action nominals form plurals?

- – can they be plural under the process/event reading (as opposed to the object/result reading)? if so, do they pattern like other nouns? any interaction of plural marking with aspect? if necessary, comments on their semantics
- nominalized adjectives
  - if the plurals of headless attributes are identical to their form when used with plural heads, simply state so;
  - nominalized attributes may acquire nominal declension, including plural marking – especially typical of highly agglutinative languages (e.g. Altaic);
  - when headless, attributes may acquire special morphological properties; if this bears on their plural, describe the relevant phenomena.

2.4 Verbal number

- is there evidence for verbal number as distinct from agreement? if so, what criteria make it distinct from simple agreement with a plural argument?
  - e.g., special verb forms for expressing distinct events, e.g. Hausa *Yaa sÀs-sÀyɪ abuubuwaa* 'he BOUGHT many things' (many different types of things, or in many events); *naa ZUZ ZÙBÀ shaayì* 'he POURED tea (for different people, not repeatedly)' (Součková 2011)?
- how widespread is it in the lexicon?
  - are there common singular-plural pairs of *formally suppletive* verbs, or only marginal *plural-argument-only* verbs (like *massacre*)?
- is there evidence that the morphology expressing multiple-event readings is distinct from the morphology of verbal aspect?
  - e.g., any evidence that numerosity-of-events category differs from a generic iterative / habitual imperfective specification? what types of distributivity other than distributivity over time does it express?
- how does verbal number differ from (pro)nominal number?
  - same values, same exponence? special verbal categories (e.g. *collective*, to explicate)? is suppletion common for verbs expressing multiple participants and/or events? if so, provide a (nearly) exhaustive list
- numerosity or multiplicity of events?
  - does verbal number convey 'many' or 'more than one' meaning? is the means expressing verbal number (e.g., pluractional morpheme, or the suppletive *pluractional stem*) felicitous to indicate two occurrences?
- verbal duals: is there a special marker for two events?
  - if present, is verbal dual different from *refactive* marking that designates a second occurrence of an event (rather than two occurrences of an event)?
- is the non-plural number form semantically unmarked?
  - i.e. can a verbal form that does not overtly express event non-singularity refer to several events, or is a specifically marked form obligatory under these circumstances?

3 Agreement and the syntax of number

- number agreement in NP domain
    - describe the morphology of agreement in number on adjectives and other attributes (including demonstratives, possessive pronouns, numerals)
    - special mention if agreement is limited to a subclass of adjectives, or otherwise divergent behaviour of different classes of adjectives, like a few forming plurals irregularly (such as stem suppletion for number); or adjectives showing special type of syncretisms and cumulation in their declension
    - special mention if, in terms of their morphosyntactic behaviour, property words are attributivized predicates rather than primary attributes
- number agreement inside DP/NP domain and/or inside the clause; targets, domains and conditions
    - what can or must agree in number with the head noun (determiners, modifiers; pre- or post-head asymmetries) or with the controller NP (predicate, adverbs, anything else)?
    - does DP-agreement in number with V (or other targets, like predicative adjectives) differ in any way from DP-internal agreement?
    - do complementizers (conjunctions similar to 'that', wh-words) agree in number? under what conditions (e.g. linear adjacency)?
- patterns of numeral modification
    - obligatory / optional / impossible exponence of number on a noun modified by a numeral?
    - do all numerals behave uniformly? do all nouns behave uniformly? do changes in syntactic structure affect agreement? (e.g., linear order of numeral and noun, or intervening material)
    - if there is a dual, does it necessarily appear on the noun with 'two'? may it appear on the noun with 'two' at all?
    - if there is a paucal, does it necessarily appear on the noun with 'a few'? may it appear on the noun with 'a few' what happens when numerals are coordinated (as in 'three or four N')?
    - are there unexpected choices of number on nouns modified by e.g. 1.5 or 0.5?
    - how does number interact with the system of quantifiers? do determiners/ quantifiers denoting pluralities impose a number value that would otherwise only denote singularities or masses (e.g. Irish *cúpla lá* 'a few day[sg]')?
    - does the numeral itself agrees with the noun in number or other categories?
    - can a bare numeral (i.e. not attached to a noun) act as predicate as in *the apostles were twelve* ?
    - does numeral modification affect the agreement between a NP/DP and a verb? (especially when a subject denoting a plurality is formally singular, as in 'six cow[sg]' … eat/eats).

- word order with numeral modification
  - do numerals differ from other modifiers? are there possible alternative orders?
  - are there any semantic contrasts associated with alternating orders? (for instance, in Russian postposed numerals express approximate cardinality)
- case patterns specific to numeral modification
  - do numeral govern certain choices of case, or special forms of the noun? (e.g. in Russian)
- variable loci for number marking according to the syntactic structure of NP/DP
  - plural sometimes marked on both N and Det, sometimes only on N, sometimes only on Det, depending on syntactic structure;
  - does number get realized in non-canonical positions / on non-canonical targets (otherwise not number inflected) in some contexts? (like the French plural [-z] which can surface on an adverb in *Tu fais quoi, comme[-z] études?*)
- number agreement under coordination
  - number agreement inside DP, e.g. [Det [N sg and N pl (modifiers)]]
  - number agreement when two noun phrases with different number values are coordinated
  - resolution rules for agreement conflict; cases like *this/*these man and woman*.
- number agreement on verbs
  - describe the morphology of number indexing / agreement on the verb (predicate), including when expressed syncretically with person;
  - special attention to the exponence of number in third person indexes (singular vs. non-singular);
  - special mention if number indexing is morphologically independent from agreement in person, e.g. if there is no person agreement (e.g. Nivkh, East Caucasian), or if some plural indexes show formal affinity to nominal number (third person plural in Turkic);
  - special mention of mismatches between the number as expressed on the noun and the number agreement with / indexing of the same NP on the verb (e.g. semantic agreement and related phenomena);
  - if agreement in the clausal domain goes beyond indexing on the verb (agreeing adverbs etc.), describe the morphology and conditions of their agreement.
- number categories in imperatives and commands
  - is expression of number and person or number alone on commands similar or distinct from indicative forms, in terms of values and exponence? what are the functional types of command involved?
  - can it be identified as expression of the number of addressee distinct from other means of indexing? (e.g. S/A agreement in commands in a language with an otherwise ergative agreement)

4 Semantics and Discourse

- pragmatic functions of number marking
  - primarily, politeness in pronouns (including honorific uses of inclusive, second person plural, third person plural); other relevant evidence;
  - briefly, what other items are used in this function if not plural pronouns (e.g. nouns).
- genericity in pronouns
  - is the plural or any other form used (also indicating the person of the pronoun used as generic – cf. *you never know*)?
  - briefly, what other items are used in this function if not plural pronouns? (e.g. singular pronouns, dedicated lexical items etc.)
- number (especially plural) in generic sentences
  - number values primarily associated with genericity (e.g. *beavers build dams* – cf. *each beaver builds dams*) and in reference to kinds (*bears are extinct* – cf. **each bear is extinct*);
  - in generic contexts, must the noun be non-singular, or can it only be singular, or both? what are the relevant factors, if known at all?
  - can there be a determiner, e.g. *the bears are extinct*? what difference does it make?
  - what happens to mass nouns and other uncountable nouns in such contexts?
- countability
  - are plural nouns all count (i.e., do they head count NPs)? or can plural nouns have the distribution of mass nouns? (i.e. can plural nouns take mass determiners, disallow count determiners, disallow reference to individuals by reciprocals, disallow numeral modification etc.; e.g. *the waters of the lake* but **a few waters*, **three waters*, **the waters meet each other*.)
  - *Note*: in languages where certain morphological patterns are directly related to countability, like in special number marking or number agreement strategies for mass or count nouns, or extensive alignment of singularia or pluralia tantum with a mass reading, the information in this section may be cross-referenced in Section 2.3, or re-partitioned between 4 and 2.3.
- non-plural reference of plural forms (*non-strict* reading of plurals)
  - in systems where each form has its own number value (so no number value is *general*), are there particular contexts where e.g. a plural admits non-plural reference, or a singular a non-singular one?
  - *Note*: Typical distinguishing cases are negative and interrogative sentences, like *did the boys arrive?* or *no boys arrived*: in English interrogatives, one can use the plural *boys* even if the question can be answered truthfully *yes, exactly one boy arrived*; and the negative sentence is false if exactly one boy arrived, but not more.

- – are there any additional discourse conditions for *non-strict* reading of plurals?
- relatedly, is there evidence that singular or plural presuppose (are only defined for) a single or multiple denotation?
  - – e.g., Spanish distinguishes *quién fue a la festa?* 'who[sg] was at the party? from *quiénes fueron a la festa?* 'who[pl] was at the party?', and the singular *quién* is compatible with one or many being at the party, but the plural cannot be used if it is known that only one person was at the party.
- any other known discourse-related functions of number
  - – as one possible example, is number used for tracking discourse referents (singular and plural marking exploited to track different referents)?

5 Conclusions

- What are the most notable (typologically unusual, unexpected, informative) aspects of the number category in the language?
- What is the best way to characterize the meaning of the number feature and its values in this language? How close is the fit between linguistic number marking and its core conceptual content (numerosity)?
- Does number in this language display a roughly uniform morphological and semantic organization across all word classes it applies to, or are there sharp formal differences, for instance, between nominal and pronominal number?
- Are there any clear indications of how number as a category is relevant to discourse structure?
- What are the areas of the grammar of the language related or relevant to number that are not well understood and would benefit from further research?

References

Broschart, Jürgen. 2000. Isolation of units and unification of isolates: the gestalt-functions of classifiers. in Gunter Senft (ed.), Systems of Nominal Classification, 239–269. Cambridge: Cambridge University Press.

Corbett, Greville. 2000. Number. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Ó Siadhail, Mícheál. 1989. Modern Irish. Grammatical Structure and Dialectal Variation. Cambridge: Cambridge University Press.

Součková, Katerina. 2011. Pluractionality in Hausa. Netherlands Graduate School of Linguistics.

Zhang, Niina. 2012. Countability and numeral classifiers in Mandarin Chinese. In Diane Massam (ed.), Count and Mass Across Languages, 220–237. Oxford: Oxford University Press.

Africa

Valentin Vydrin

2 Number in Eastern Dan

Abstract: Number in Eastern Dan is marked, by default, by a plural word, *dhṳ̀n*. There are also some other formal means: a specialized plural suffix for kinship terms and agent nouns; reduplication and tonal modifications for adjectives. The irregularities in plural marking are explained, to a great extent, by the Animacy Hierarchy. There are formal splits, marked in different ways:

– between the 1st person and other pronouns;
– between the pronouns for interlocutors and the 3rd person;
– between the agent nouns and some kinship terms, on one hand, and all the other nouns, on the other;
– between the names for humans, big animals and inanimate objects.

There are two mechanisms to express verbal number. One of them, the full reduplication of the verb, is typologically common, while the other one, pluralization of the preverb, is unusual. Neither of these mechanisms can be regarded as an agreement with arguments; these are cases of a true pluriactionality.

1 Overview

1.1 Information about the language

Eastern Dan is a group of dialects of the Dan macrolanguage (South Mande < Southeastern Mande < Mande < Niger-Congo) spoken in Western Côte d'Ivoire by some 650,000 people; the total number of Dan speakers (in Côte d'Ivoire, Liberia and, marginally, in Guinea) is more than 1,600,000. In this study, Eastern Dan is represented by the Gwɛɛtaa dialect, which was selected in the 1970s as the standard variety for the written norm of Eastern Dan. The data for this study were collected by the author in the course of regular field trips to Côte d'Ivoire carried out since 2001. Additionally, some texts taken from written literature were analyzed.

Eastern Dan is, basically, an isolating language, with some important elements of fusion, most of which are of a relatively recent origin: case system in "declinable nouns" (a subclass of nouns); series of personal pronouns and "predicative markers" (auxiliary words expressing TAM and polarity meanings); tonal inflection on verbs. Eastern Dan has no noun classes/genders; noun classifiers are marginal.

Eastern Dan is a five-level tone language (xH for extra-high shown as *ắ*; H for high shown as *á*; M for mid shown as *ā*; L for low shown as *à*; xL for extra-low transcribed as *ạ̀*), there are also three falling tones: H-xL for high-extralow shown

as *â*; xH-xL for extrahigh-extralow (very rare) shown as *ä*; M-xL for mid-extralow (very rare) designated as *ã*.[1]

The basic word order is rigid: in a clause with a verbal predicate it is (S) Aux (O) V (X).[2] The subject is indexed on the Aux and is often absent; presence of a direct object in a transitive verbal construction is obligatory, its absence renders the verbal construction intransitive, and the verb acquires a decausative or passive meaning. In an NP, the dependent noun precedes the head noun, and the dependent adjective follows the head noun. Determiners follow the noun; if a noun is followed by an adjective, the determiners follow the adjective (some exceptions will be discussed in this article). Therefore, the structure of an NP in Eastern Dan is as follows: [[NP] N [Adj] Det].

The class of nouns is subdivided into two subclasses: "inflexible nouns" (further on, just "nouns"), and "declinable nouns", see in detail (Vydrin 2011). The latter subclass is characterized by the existence of noun declension (otherwise highly atypical of Mande languages, and of Niger-Congo languages in general). Among the "declinable nouns" there are frequent words for body parts and place names (such as 'village', 'farm', 'land', etc.). The declension system of the declinable nouns is irregular; most of them (place names, including toponyms) have two cases (common and locative), others may have up to four cases. The particular property of the declinable nouns is that they are used in the oblique function without postpositions (unlike "true nouns"), which is explained by their originating from fusion of nouns with postpositions.

1.2 Overview of the number category

The number in Eastern Dan is a quasi-obligatory and quasi-regular category. The major plural marker is the plural word *dhừn*. The use of *dhừn* with nouns depends on their position in the Animacy Hierarchy. In addition, at the top of the Animacy Hierarchy one observes some suppletion and an unproductive nominal plural suffix *zʌ̀*. Adjectives have irregular models of plural-intensive formation.

In the pronominal system (both in personal pronouns and predicative markers inflected for person and number), Dan has dual inclusive forms distinct from the plural inclusive ones.

---

1 This paper uses the 2014 orthography for Eastern Dan. In this orthography, absence of a tonal mark on a vowel means that the tone is identical to the preceding tone (e.g., *bāa* /bāā/ 'cassava'); *dh* and *bh* stand for implosive consonants /ɗ/ and /ɓ/; vocalic nasality is designated by the letter *n* following the vowels of a foot (e.g., *kwàan* for /kwȁȁ/ 'to steal'); *ɯ, ɤ, ʌ* stand for back unrounded vowels; *œ* is an open labialized vowel, and *æ* is a front open vowel.

2 S for the subject, O for the direct object, X for an oblique (adverbial, postpositional phrase, etc.), V for the verbal predicate, Aux for the predicative marker or bifunctional auxiliary (i.e. an auxiliary word which can appear both as a predicative marker in a verbal clause or as a copula in a non-verbal clause).

There are no widespread alternative plural forms for the same item (only some isolated cases, hardly differentiated semantically).

The verbal number is expressed by the reduplication of verbs or by the plural word *dhɯ̀n* inserted between the preverb and the verbal stem.

The only case of clausal agreement in number is that of predicative markers with the subject NP. At the NP-internal level, there are instances of expression of number through modification of the stem of an adjective. Another phenomenon that may be interpreted as NP-internal agreement is the optional appearance of the plural word *dhɯ̀n* both after a noun and an adjective (by default, it appears only after the adjective, i.e. at the end of the NP).

2 Pronominal, nominal and verbal number

2.1 Generalities

With respect to number, nouns, pronouns and adjectives share one and the same basic plural marker, a plural word *dhɯ̀n*. Pronouns have the third value of the number category, the dual. Adjectives have a specific mechanism of derivation of plural forms, cumulatively with intensity, through reduplication, tonal modification and transfixation.

Verbal number is expressed by full reduplication of verbal stems or by the plural word *dhɯ̀n* inserted between a preverb and the verbal stem. These mechanisms are different from agreement.

2.2 Pronominal number

2.2.1 Personal pronouns

Personal pronouns in Eastern Dan have a number of series which are differentiated by their syntactic and pragmatic functions, see Table 1 (inclusory pronouns will be analyzed separately, in §2.2.3).

As can be seen, pronouns have singular, dual and plural forms.[3]

---

3 The first person inclusive dual pronoun in systems of the same type as Dan is usually considered as belonging to the singular part of the pronominal paradigm (i.e. to the "restricted" pronouns, vs. "non-restricted", or plural), cf. in particular (Greenberg 1988). I do not contest this interpretation; however, we should not discard the fact that the dual (1st+2nd) pronoun refers to two persons and therefore stands out against the other restricted/singular pronouns which have reference to only one person or object. Cf. (Corbett 2000: 91) for an interpretation of the 1st pers. dual pronoun within

Tab. 1: Eastern Dan personal pronouns.

Person	Singular				Dual	Plural					
	1	**2**	**3**	**Refl.**	**Incl.**	**1 inclusive**	**1 exclusive**	**2**	**3**	**Refl.**	
Non-subject	ŋ̄	ī/ū	à̰	ɤ̄	kō	kwā	yī	kā	à̰-dhʉ̀n	wō	
Possessive	bhān	ū bhā	à̰ bhà̰	ɤ̄ bhā	kōò̰	kwāà̰	yīḭ̀/yāà̰	kāà̰	à̰-dhà̰an	wōò̰	
Possessive honorific						kwēɛ́		kēɛ́			
Autonomous	bhān	bhī	yɤ̄	–	kō	kwā	yī	kā	wō	–	
Restrictive	bhānŋ/bhāan	bhīi	yɤ̄ɤ	–	kōo	kwāa	yīi	kāa	wōo	–	
Selective	bhán	bhí	yɤ́		kó	kwá	yí	ká	wó		
Negative focalized	bhāáan	bhīáa	yɤ̄áa		kōáa	kwāáa	yīyáa	kāáa	wōáa		

Dual is distinguished only in the first person inclusive, represented by the segmental base KO, while the plural inclusive pronouns have the base KWA. Etymologically, KWA goes back to the combination of first person inclusive (what is now dual inclusive) and the second plural forms, presumably *koka. Most of the plural pronoun forms are suppletive with respect to their singular counterparts. There is however an important exception, predictable on typological grounds (Corbett 2000: 62–64): in the non-subject series, the third plural form, *à̰-dhʉ̀n*, is derived from the third singular form, *à̰*, with the means of the regular plural word *dhʉ̀n*. The possessive third plural pronoun *à̰-dhà̰an* results from fusion of the non-subject third plural pronoun *à̰-dhʉ̀n* with the possessive marker *bhà̰*.

Reflexive pronouns constitute a suppletive opposition in number. Formally, they are similar to personal pronouns and can be regarded as belonging to the same paradigm.

the logic of the Animacy Hierarchy. Besides, note that the fact that the plural inclusive may also refer to the speaker and the addressee – i.e. to only two people, see below examples (1) and (2) – is not fully compatible with the logic of the minimal-augmented systems.

2.2.2 Predicative markers

There are no clitic pronouns in Eastern Dan. There are auxiliary words, referred to in the Mandeist linguistic tradition as "predicative markers", which are syntactic heads of clauses; they are inflected for person and number, TAM and polarity, see Table 2.

As we can see, the predicative markers do not differ very much from the personal pronouns with respect to the expression of number, with one notable exception: unlike in non-subject pronouns, the number in all persons including third person is suppletive. The opposition of the dual and plural in PMs is present in first person inclusive. Logophoric PMs are suppletive in number in the same way as the reflexive pronouns.

The opposition between the dual and the plural in predicative markers is optional, and dual forms may be replaced by plural ones. Cf. (1), where a dual PM is

Tab. 2: Eastern Dan predicative markers.

| | Singular | | | | Dual | Plural | | | | | |
Person	1	2	3	Log.	Incl.	1 inclusive	1 exclusive	2	3	Log.
Existential	ā	ī/ū	yɤ̀	ɤ̄	kō	kwā	yī	kā	wɔ̀	wō
Conjoint	á	í/ú	Ø/ɤ́/yɤ́	ɤ́	kó	kwá	yí	ká	wó	wó
Perfect	bhán	bhá	yà/yǎ	yá	kó	kwá	yá	ká	wà/wǎ	wá
Prospective	bhāan	bhīi/yīi	yɤ̄ɤ	–	kōo	kwāa	yīi	kāa	wōo	–
Imperative	–	Ø/bhɤ̀	–	–	kɔ̀	kwà	–	kà	–	–
Subjunctive	á	í/ú	Ø/yɤ̀	ɤ̌	kó	kwá	yí	ká	wɔ̀	wɤ́
Presumptive	bhāàn	bhāà	yāà	–	kōɔ̀	kwāà	yāà	kāà	wāà	–
Negative imperfective	bháan	bháa	yáa/áa	–	kóo	kwáa	yáa	káa	wáa	–
Negative gnomic			áan							
Negative perfective	bhíin	bhíi	yíi	–	kóo	kwíi	yíi	kíi	wíi	–
Negative subjunctive	bhán	bhá	yá	–	kó	kwá	yá	ká	wá	–
Negative consecutive	bhín	bhí	yí	–		kwí	yí	kí	wí	–

used, and (2), where a plural one appears, although only two participants are referred to (the context of the narrative leaves no doubt that only two participants, the speaker and his interlocutor, are concerned).

(1) *Kō   tàabhǎn!*
1DU.INCL.EXI behind
'See you again!' [f063:18], lit: 'We (two) are after' ("we two will be together next time").

(2) ***Kwǎ*** *zlɔ̀ɔ* *yɛ́ɛdhɤ̄* ***kwá*** *dhó* ***kwā*** *tàa*
1PL.INCL.IMP slow.down slowly 1PL.INCL.SBJV go 1PL.INCL.NSBJ back.LOC
kɤ̄ ***kwá*** *kɑ̄ dhɛ̀ bhɛ́n* *dhɤ́ ɤ́* *dùa-sīʌ pɑ̄ gɔ̀.*
SMLT 1PL.INCL.NSBJ do that human/REL be 3SG.JNT flee-DUR thing PP
'Let us go slowly behind, as someone who is trying to avoid something' [Wɔ́n dhɤ́].

Optionality of the dual in Eastern Dan thus follows the pattern described in (Corbett 2000: 42ff).

2.2.3 Inclusory and compound pronouns

Eastern Dan has inclusory constructions and a dedicated series of personal inclusory pronouns (for an analysis of inclusory pronouns in Mande languages, see (Vydrin 2010; Khachaturyan 2019)). In fact, these constructions are the default means of coordination in this language.[4]

There are two series of inclusory pronouns, see Table 3.

Tab. 3: Inclusory and coordinative compound pronouns.

Person	1 inclusive	1 exclusive	2	3
Inclusory simple	*kwāa ~ kwāǎ*	*yāa ~ yāǎ*	*kāa ~ kāǎ*	*wāa ~ wāǎ*
Inclusory bound	*kwāǎ*	*yāǎ*	*kāǎ*	*wāǎ*
Coordinative compounds	*kwēŋ ~ kwḕŋ̀*	*yḕŋ̀*	*kēŋ ~ kḕŋ̀*	*wēŋ ~ wḕŋ̀*

It is obvious that these pronouns go back to the respective plural personal pronouns. In these series, quite naturally, there is no opposition of dual and plural forms.

---

4 There is also a coordinative conjunction *bhān ~ bhàn,* but it is rarely used and can be regarded as a marginal means of coordination in Eastern Dan.

The pronouns of the **simple inclusory series** appear as the first component of a coordinative construction, the other conjunct being represented by an NP.

(3) **Yā̀à** Gbằtờ yá dhūn.
    1.EXCL.ICL NOM.M 1PL.EXCL.PRF come
    'I and Gbato, we have come' (lit.: 'We and Gbato …').

If the first conjunct is also expressed by a non-pronominal NP, the third person simple inclusory pronoun is inserted between both NPs (thus gaining some similarity with a coordinative conjunction):

(4) Sī̋tằ **wāà** Dá̋ằn dhú wờ wō kó gɔ̋-lɛ̀
    NOM.F 3ICL NOM.M daughter 3PL.EXI PL.REFL RECP.CMM head-tie\NEUT
    dhɔ̋ɔzīandō ká.
    Saturday with
    'Sita and Dan's daughter plait each other's hair on Saturdays' (lit.: 'Sita, they and Dan …').

The covert subset of the inclusory pronouns in (3) and (4) consists of one member (in Beam de Azcona's terms (2006), cit. by (Khachaturyan 2019)): (3) has no reading *'we and Gbato'. If one of the conjuncts needs to be pluralized, the plural word -*dh̀ùn* appears after the second conjunct. However, this procedure allows for ambiguity, because the scope of the plural word is unclear: it may be the first conjunct, the second conjunct, or both. So, in (5), three interpretations are possible.

(5) Yāà dhēbɛ̀-dhʌ̀ùn yá dhūn.
    1.EXCL.ICL woman-PL 1PL.EXCL.PRF come
    'We and a woman', or 'I and women' or 'We and women'.

In other words, the scope of the plural marker is the entire coordinated construction.

The inclusory pronouns serve for coordination of any type of nouns, whatever may be their position in the Animacy Hierarchy. See (6) where inanimate nouns are coordinated.

(6) Dhēbɛ̀dhʌ́n bhā yằ bāa wāà slʌ̀ dhɔ́.
    woman ART 3SG.PRF cassava 3ICL pepper buy
    'The woman has bought cassava and pepper' [f056:05].

The pronouns of the **bound inclusory series** differ from those of the simple inclusory series formally by the lack of variability of the final tone. They are used when the second conjunct includes, at the underlying level, a non-subject 3SG pronoun *ằ*

(according to the word order within the NP, it occupies the left-most position). This pronoun is incorporated into the inclusory pronoun:

(7) *Yāằ*   *gbɤ̄ yá*   *dhūn.*
1EXCL.ICL>3SG.NSBJ son 1PL.EXCL.PRF come
'I and his son, we have come'.

The bound series is derived from the simple series through fusion.

    The **coordinative compound pronouns** are used to express coordination of two pronouns. The 1st inclusive compound pronoun (*kwɛ̄ŋ ~ kwɛ̄ŋ̈*) refers to the 1st+2nd persons (I and you sg.), and for all the others, the second conjunct is third person (I and him; you sg. and him; he and him).

    The compound pronoun cannot refer to more than two participants. In (8), these are the speaker and the subprefect.

(8) *Kűnbhằndhằn sɛ̄ɛ́ndhʌ̀n yɤ̀*   * n̄*   *dhɛ̀*   *kɤ̄*   **yɛ̄ŋ**
prefect   small   3SG.EXI 1SG.NSBJ call\NEUT SMLT 1SG.3SG
yí   *pʌ̄*   *bhɤ̀ yī*   *kwʌ́ʌ̀.*
1PL.EXCL.SBJV thing eat 1PL.EXCL.NSBJ RECP.SUP
'The subprefect invited me to eat together' (lit.: '... so that we two eat together') [e16:17].

If the reference is made to more than two participants, the compound pronoun is followed by the plural word *dhʌ̀n*. Like with the simple inclusory pronouns, the scope of the plural word is ambiguous. At the underlying level, two conjuncts are meant, but it is not specified which of them is pluralized: the first, the second, or both, see (9).

(9) *Wɛ̄ŋ-dhʌ̀n wɔ̀*   *dhɔ̀*   *dhó'.*
3&3-PL   3PL.EXI go\NEUT go\INF
'They (3PL&3SG, or 3SG&3PL, or 3PL&3PL) will go'.

2.2.4 Interrogative pronouns

The interrogative pronouns *dē* 'who?' (about a human) and *bhʌ̀n* 'what?' (about inanimate objects and animals) can be followed by the plural word *dhʌ̀n*. This implies that more than one human (10a) or object (11a) is meant, and these humans or objects are perceived as more or less homogeneous groups.

(10) a. *Dē-dhʌ̀n ɤ́*   *bhá*   *ằ-dhʌ̀n*   *dhɛ̀?*
      who-PL   REL 2SG.PRF 3SG.NSBJ-PL call
      'Which persons have you called?' (a natural answer: "Soldiers" or "Hunters").

(11) a. *Zân   yà̄   bhɰ̀n-dhɰ̀n dhɔ́?*
      NOM.M 3SG.PRF what-PL   buy
      'Which things has Zan bought?' (a natural answer: "Hoes" or "Bikes").

Another strategy of pluralization is observed in the construction where two identical interrogative pronouns are coordinated. In this case the speaker means that there are two people or objects in question, and he expects them to be enumerated; this is a distributive dual interrogative construction.

(10) b. *Dē   wāà̄ dē   ɤ́   bhá   à̄-dhɰ̀n   dhɛ̀?*
      who 3ICL who REL 2SG.PRF 3SG.NSBJ-PL call
      'Whom (dual) have you called?' (a possible answer: "Tokpa and Yo").

(11) b. *Zân   yà̄   bhɰ̀n wāà̄ bhɰ̀n dhɔ́?*
      NOM.M 3SG.PRF what 3ICL what buy
      'What (dual) has Zan bought?' (a possible answer: "A hoe and a knife").

Finally, interrogative pronouns can be coordinated, and the second one can be followed by the plural word *dhɰ̀n*. This is a distributive plural interrogative construction.

(10) c. *Bhá   dē   wāà̄ dē-dhɰ̀n dhɛ̀?*
      2SG.PRF who 3ICL who-PL call
      'Whom (pl.) have called?' (a possible answer: "Tokpa, Gbato, Sayi and Yo").

(11) c. *Ī̄   glà̄   bhɰ̀n wāà̄ bhɰ̀n-dhɰ̀n dhɰ̀n gú?*
      2SG.EXI transform/NEUT what 3ICL what-PL   FOC in
      'You transform yourself into what?' [Wɔ́n dhɤ́] (the context is as follows:
      a buffalo transformed into a young woman tries to get the secret of magic
      transformations out of the great hunter; obviously, the buffalo-woman
      wants to know all his reincarnations).

2.3 Nominal number

2.3.1 Typological profile

The main means of expression of nominal plurality is a plural word *dhɰ̀n*. There is a small group of names of humans whose pluralization implies suppletion or substitution of the agent suffix. Classifiers are used with some nouns, expressing the singulative meaning or appearing in constructions with numerals.

The optionality of the plural word is sensitive to the Animacy Hierarchy. Irregular plural forms are only observed for humans.

Nouns have no dual.

2.3.2 Exponence

No cumulation of the plural with case or definiteness is observed in Eastern Dan. In adjectives, there is cumulation of the plural meaning with the intensity.

The standard plural marker is the plural word *dhʋ̀n* (I classify it together with determiners) which is postposed to the noun.[5]

In Eastern Dan, there is only one noun that has a suppletive plural form: sg. *dhēbhʎ̀* 'woman' – pl. *dhōo-dhʋ̀n ~ dhōŋ-dhʋ̀n* 'women' (a regular plural form *dhēbʎ̀-dhʋ̀n* also exists, and it is more frequently used than the suppletive one).[6] More precisely, the plurality is expressed twice, by the plural word and by the form of the stem. The suppletive form in question also appears in compounds, e.g.: sg. *gwʎ̄ʎ́dhē* 'maidservant' – pl. *gwʎ̄ʎ́dhōo-dhʋ̀n ~ gwʎ̄ʎ́dhōŋ-dhʋ̀n ~ gwʎ̄ʎ́dhēbhʎ̀-dhʋ̀n*.

Nouns which have in their singular form a suffix *-bhīn*[7] designating human males, have their suffix replaced by another one, *-zʎ̀*.[8] At the same time, the plural word *dhʋ̀n* is added, see Table 4.

The suffix *-zʎ̀* is also added to some names of relatives or partners in relations, younger or age-mates, often irrespective of their sex; in the singular, these nouns do not have any suffix, see Table 5.

As we can see, some of these nouns display variation in the plural formation: both forms with and without *-zʎ̀* are possible (*dèbhīn* 'soothsayer', *gwʎ̄ʎ́bhīn* 'factotum', *dhʎ̌sʎ̌* 'partner in joking relation', *tébhʎ̌n* 'age-mate', and for *ző̄obhīn* 'witch doctor', even three variants of plural are available).

---

5 In agreement with the rules of the 2014 orthography, it is connected with the preceding word by a dash.

6 The etymology of the form *dhōo ~ dhōŋ* is unclear. It may result from a contraction of the word for 'woman', originally **dēbɔ̌*, where *-bɔ* is an archaic suffix for humans (in Dan Gwɛɛtaa **-bɔ̌* > *-bʎ̀*). On the other hand, it may be a stem different from *dhēbʎ̀*, cf. two different roots in closely related languages Tura and Goo: *lé* 'wife' and *lóŏ* 'woman' (Aplonova 2020; Bearth 1971).

7 The suffix *-bhīn* surely goes back to a noun for 'human'. In modern Dan, 'human' is *bhɛ̄n*, it may refer to a man or a woman.

8 As the list shows, some nouns have alternative plural forms preserving the suffix *-bhīn*, or without any suffix altogether.

In Kla-Dan (the northernmost variety of Dan) there is a suffix *-zʎ̀* for male agents which appears in both singular and plural forms, as an equivalent of *-bhìn*, e.g.: *lùèŋ́ɓɨ̀ ~ lùèŋ́zʎ̀* 'hunter', pl. *lùèŋ́ɓɨ̀-lʋ̀ ~ lùèŋ́zʎ̀-lʋ̀* (Makeeva 2017: 657). In Western Dan, suffix *-zʎ̀* is a plural marker for three kinship terms: *zíŷzʎ̀-dhʋ̀n* 'grandfathers', *bházʎ̀* or *bhɔ̀ʋŋ̀zʎ̀-dhʋ̀n* 'age-mates' (Erman & Loh 2008: 204–205), *zláädhʎ̀dhēnzʎ̀-dhʋ̀n* 'younger sister' (Alexey Fedorenko, p.c.). As we can see, *-zʎ̀* in other Dan varieties is more restricted in its use than in Eastern Dan, but it remains in the same domain.

Tab. 4: Nouns with the suffix *-bhīn* in singular replaced by the suffix *-zʌ̀* in plural.

Singular	Plural	Meaning
bhɔ́ɔ̀nbhīn	*bhɔ́ɔ̀nzʌ̀-dhʉ̀n*	'imam'
dɛ̀bhīn	*dɛ̀zʌ̀-dhʉ̀n ~ dɛ̀bhīn-dhʉ̀n*	'soothsayer'
dhɔ̀tlɔ́ɔ̀bhīn	*dhɔ̀tlɔ́ɔ̀zʌ̀-dhʉ̀n*	'medical doctor'
dhūeebhīn	*dhūeezʌ̀-dhʉ̀n*	'hunter'
gwʌ̃́ʌ̃́bhīn	*gwʌ̃́ʌ̃́zʌ̀-dhʉ̀n ~ gwʌ̃́ʌ̃́-dhʉ̀n*	'factotum, servant; disciple' (male)
kplʌ́ʌ̀nŋbhīn	*kplʌ́ʌ̀nŋzʌ̀-dhʉ̀n*	'outstanding person'
sǽœ̀bhīn	*sǽœ̀zʌ̀-dhʉ̀n*	'archer'
yēɛbhīn	*yēɛzʌ̀-dhʉ̀n ~ yēɛgɔ̄n-dhʉ̀n*	'griot' (member of an inferior caste)
yɔ̀bhīn	*yɔ̀zʌ̀-dhʉ̀n*	'father-in-law'
zőobhīn	*zőozʌ̀-dhʉ̀n ~ zőo-dhʉ̀n ~ zőobhīn-dhʉ̀n.*	'witch doctor'

Tab. 5: Nouns which are suffixless in singular and acquire the suffix *-zʌ̀* in plural.

Singular	Meaning, sg.	Plural	Meaning, pl.
bɛ́ɛ̀	'nephew, niece' (sister's child)	*bɛ́ɛ̀zʌ̀-dhʉ̀n*	'nephews' (only males)
dhʌ̃́sʌ̃́	'partner in joking relation'	*dhʌ̃́sʌ̃́-dhʉ̀n, dhʌ̃́sʌ̃́zʌ̀-dhʉ̀n*	'partners in joking relation'
dhēébhāŋ	'sibling'	*dhēébhāŋzʌ̀-dhʉ̀n*	'siblings'
dhínbhɔ́ɔ̀n	'in-law' (male or female)	*dhínbhɔ́ɔ̀nzʌ̀-dhʉ̀n*	'in-laws' (only males)
tőbhʌ̃́n	'age-mate' (male or female)	*tőbhʌ̃́nzʌ̀-dhʉ̀n ~ tőbhʌ̃́n-dhʉ̀n*	'age-mates'
tŐœ̀ŋdʌ̄, tŐŋ̀dʌ̄	'intimate friend'	*tŐœ̀ŋdʌ̄zʌ̀-dhʉ̀n, tŐŋ̀dʌ̄zʌ̀-dhʉ̀n*	'intimate friends'

Another irregularity is observed in *tɛ̋ɛdō* 'friend, colleague'. In the plural, its final element *dō* is replaced by a non-productive human suffix *-bʌ̀* (*tɛ̋ɛbʌ̀-dhʉ̀n* 'friends').[9]

2.3.3 Numeral classifiers

Numeral classifiers in Eastern Dan go back to nouns and are at an early stage of grammaticalization. Most nouns are used with numerals without any classifiers; for others, classifiers are often optional. The situation is further complicated by the fact that the same noun roots are sometimes integrated into compound nouns. In some

---

9 The element *dō* may go back to the numeral *dō* 'one'. *-bʌ̀* is a fossilized suffix for humans, its semantics is vague.

cases, it is not clear whether the element in question should be regarded as a classifier or as a component of the compound.

Three nouns have advanced towards the status of numeral classifiers more than the others.

- *gā* 'bone, grain'. This classifier is by far the most grammaticalized in Dan (for a detailed semantic analysis of the corresponding element in Western Dan, see (Erman 2005)). It appears as a singulative marker with some names of body parts (*yǎn* 'eye(s)', *sɔ̌n* 'tooth/teeth'), names of small animals (fishes, insects), names of inanimate objects (*dhɛ̋* 'leaf', *gwʌ̀* 'stone'), etc. This classifier may be obligatory (12) or optional (13), depending on the semantics of the noun.

  (12) *Bhán   dhɛ̂ɛ̀n gā   sɔ̌ɔdhṹ dhɛ̋-kā.*
       1SG.PRF tongue BONE five   heal
       'I have healed five tongues' [g05_17].

  (13) *Bhán   vɛ̋ (gā) sɔ̌ɔdhṹ dhɔ́.*
       1SG.PRF fry BONE five   buy
       'I have bought five fries' [g05_03].

- *bhē* 'fruit'. As a classifier, it appears with names of fruits (its presence is then obligatory in numeral phrases) and of certain elongated objects (e.g. *dhɪ̀ɪaa* 'earthworm', *kpʌ̄nŋ* 'centipede', *yúr̄r̄* 'fish', *blúṳ̀* 'bread'), in which case it is optional;

  (14) *Bhán   kpʌ̀n kpʌ̄nŋ   (bhē) sɔ̌ɔdhṹ bhʌ̀.*
       1SG.PRF see   centipede FRUIT five   on
       'I have seen five centipedes' [g04_12a].

- *dhʌ́n* 'child'. As a classifier, it appears with animate nouns; it is most often optional. It is homonymous with the diminutive suffix (which also goes back to the noun for 'child'), so both readings are acceptable in many cases (15).

  (15) *Bhán   gbɪ̀ŋgbr̄   dhʌ́n yǎagā dhɔ́.*
       1SG.PRF electric.eel CHILD three   buy
       'I have bought three (live or small) electric eels' [g05_02b].

Some nouns can go with different classifiers, which may influence their interpretation. Thus, in (16), the use of the classifier *gā* or no classifier altogether is the default. The classifier *bhē*, on the other hand, implies that the fish are big, while the classifier *dhʌ́n* means that they have been bought alive.[10]

---

[10] *Yúr̄r̄-dhʌ́n* may also mean 'small fish', in this case *dhʌ́n* appears in the function of a diminutive suffix, rather than of a classifier.

(16) *Bhán   yúȓ̃ɣ gā / bhē̄ / dhʌ́n / Ø   yȁagā dhɔ́.*
    1SG.PRF fish   BONE / FRUIT / CHILD / Ø three   buy
    'I have bought three fish' [g05_08].

There are some other nouns which function as classifiers, although the degree of their grammaticalization seems to be inferior to that of *gā, bhē̄* and *dhʌ́n*, and they can still be interpreted as nouns.

    kpȳ̃ 'ball, piece, group' is used with names of ball-shaped objects and as a quantum of natural segmentation, e.g. *sāȁ kpȳ̃* 'cake of soap'.[11] *Kpȳ̃* can be also used as a collective classifier: *bhȁan kpȳ̃* 'flock of birds';

    gèe 'corpse, (dead) body', for carcasses of animals;

    bȳ̃ɣ 'bunch' (of fruits), for carcasses of animals.

    There is also a mensural classifier *yʌ́n*. The original meaning of this word is 'eye', as a mensural classifier it is used for the content of a container. So, the phrase (17) designates a quantity of water sufficient to fill three gourds (even if the water is not actually in the gourds). If *yʌ́n* is omitted, the resulting phrase (*yí̋ kɔ̏ɔ yȁagā*) designates exactly three gourds full of water.

(17) *yí̋   kɔ̏ɔ   yʌ́n yȁagā*
    water gourd EYE three
    'three gourds of water'

2.3.4 Interpretation of plural

Plural of mass nouns (expressed by the standard plural word *dhʉ̏n*) may have three readings, depending on the context: abundance (large quantity of X), see (18); distributivity (X in more than one container; more than one portion of X; some X here and there), see (19); sorts (20).

(18) *Ȁ   zū̄   gbȁan   yí̋   bhō̄,   ȁ   **wēé̋-dhʉ̏n** ōo,*
    3SG.NSBJ bottom all.3SG.PRF water remove 3SG.NSBJ urine-PL   any
    *ȁ   **gbō̄-dhʉ̏n** ōo, ȁ   gbȁn yà   ȁ   blȅ̈ɛ   pā.*
    3SG.NSBJ excrement-PL any 3SG.NSBJ all   3SG.PRF 3SG.NSBJ underpants fill
    'His behind grew completely wet, his urine, his excrements, all this stuff
    filled his underpants' [Wɔ́n dhɣ́].

(19) ***Yí̋-dhʉ̏n*** *wó   bhā,   dhūn ȁ-dhʉ̏n   ká.*
    water-PL 3PL.JNT there come 3SG.NSBJ-PL with
    'This water (in a number of containers), bring it' [f180:4].

---

11 Traditionally, a cake of soap in West Africa is ball-shaped. However, *sāȁ kpȳ̃* can refer to a cake of soap of any form, even not spheric.

(20) **Yǐ-dhὺn** wó   bhā, wáa   kὰan kwʌ́ʌ pʌ̀   ká.
water-PL 3PL.JNT there 3PL.NEG.IPFV mix\NMLZ RECP.COM thing\IZF with
'These sorts of water are not to be mixed' [f180:3].

Plural of nouns for proper names of humans and kinship terms can have an associative reading (21), and for non-humans, a similative reading (22), cf. (Daniel & Moravcsik 2013; Mauri & Sansò 2018: 21–23).

(21) À̰   **dhǎn-dhὺn**   gɔ̀   yǐ   bhὰa-sὺ̀ ɔ̀o, kɤ̄'
3SG.NSBJ mother.in.law-PL leave\NMLZ water by-GER EMPT SMLT\3SG.NSBJ
tɤ̌ɤdō bhā yà   lòo'   gú gbɤ̌ ká.
friend ART 3SG.PRF arrive\3SG.NSBJ in crying with
'As soon as his mother-in-law and the other one (= her daughter) left for water, his friend started crying' [Wɔ́n dhɤ́].

(22) ... yɤ́   dhʌ́nglɔ̄ɔ́ndhʌn bhā yà   glɤ̌   **dhɤ́ŋkpāadhɤ̀-dhὺn** gú
CONS youth   ART 3SG.PRF transform butterfly-PL   in
yɤ́   wà   ὰ   bhɔɔ̄̀n wáa'   yɤ̄.
CONS 3PL.PRF 3SG.NSBJ search 3PL.NEG.IPFV\3SG.NSBJ see
'... then the young man transformed himself into a butterfly and the other beings, they looked for him and did not see him' [Wɔ́n dhɤ́] (the context: the hunter had an ability to avoid his enemies through transforming himself into different objects and insects, among them, a butterfly).

There are no special forms for collectives. A collective meaning is usually expressed by a bare noun (without the plural word).

I have no information about the asymmetry between "strict" and "non-strict" readings for plural nouns.

2.3.5 Distribution across the lexicon

The availability and the use of the plural word *dhὺn* depends on the position of the referent on the Animacy Hierarchy. If a noun designates a human being, the use of the plural marker is quasi-obligatory (23). If it refers to an animal (at least, a big one, informally, endowed with individuality), it is optional, but also preferable (24). With the names of small animals (lacking individuality) or inanimate objects, the plural word does not usually appear (25, 26).

(23) **Dhʌ́n-dhὺn** wɔ̀   tlōo kʌ̄-sʌ̄ʌ kpɤ́ŋdhɤ̄.
child-PL   3PL.EXI play do-DUR outside.LOC
'Children are playing outside' [b183:02].

(24) *yɤ́ wó dǎ' ká* **sòo-dhṳ̀n** *tằ, yɤ́'   gbàn*
CONS 3SG.JNT go.up\JNT\3SG.NSBJ with horse-PL on CONS\3SG.NSBJ all
yɤ̀   yɨ̄n.
3SG.SBJV end
'And they passed to horses, and then they (horses) were over' [Wɔ́n dhɤ́].

(25) **Zɛ̄nɤ́ŋ** *yɤ̀ wě-sīʌ.*
mosquito 3SG.EXI speak-DUR
'A mosquito pipes / Mosquitos pipe.' [f079]

(26) *Kpằn ằ   bhǎ-sṳ̀ dhūn ɤ́ gbë̋ë̀ kɤ̄ wằ   wèŋ'*
see 3SG.NSBJ on-GER FOC 2SG.JNT difficult SMLT 3PL.PRF pour\3SG.NSBJ
*gú, wằ   yằ ằ   **gā** bhɤ̀-sṳ̀ bhǎ.*
in 3PL.PRF put 3SG.NSBJ grain eat-GER on
'As soon as they had seen it (the oil palm), they rushed on it and began eating its grains' [Wáa wɔ́n].

However, these are just trends, not strict rules. In natural texts, human nouns can be found without a plural marker in the contexts where their plural reference is doubtless, while names of inanimate objects may appear with this marker. Referential status certainly plays a role: non-referent human nouns, like in (27), tend to appear without the plural word, and referent inanimate nouns can be accompanied with the plural word (28).

(27) *yɤ́' dɨ̄ yɤ̀ **bhɛ̄n** dằ   yī dō ká kɤ̄*
CONS\3SG.NSBJ father 3SG.EXI human go.up\NEUT day certain with SMLT
wɔ̀ dhó' dhɛ̄ë́-kpɔ́-dhɛ̀   gú...
3PL.SBJV go\3SG.NSBJ question-spread-MSD in
'Then his father sent some people, one day, to go and to ask him...' [Wɔ́n dhɤ́].

(28) *Dhɛ̀kpɔ̀œœyɤ̀ dō ká, yɤ́   gūn zīɤ-sīʌ bṳ́ gú pɤ̀-dhɛ̀*
day one with CONS.3SG.CNJ be.PST walk-DUR bush in village-CMM
sɔ́ɔ, yɤ́   kpằn bhánŋglōo dhṳ̀ dō bhǎ kɤ̄ ằ
near, CONS.2SG.CNJ see mango tree\IZF one on SMLT 3SG.NSBJ
bhɛ̄-dhṳ̀n *wằ pɨ̄-kɨ̄.*
fruit-PL 3PL.PRF ripen
'One day he was walking in the bush near the village, and he saw a mango tree, its fruits were ripe'.

Inanimate nouns with the plural word may express the meaning of abundance (29).

(29) *Dhű̋ **kwēɛ́-dhừn** wɔ̀ yɛ̋-sīʌ, **pēŋɔ̄n-dhừn** wɔ̀ gó-sīʌ*
   tree branch-PL 3PL.EXI break-DUR flame-PL 3SG.EXI leave-DUR
ằ dhı̋ ...
3SG.NSBJ mouth
'Branches of trees were breaking, flames were shooting out of its mouth ...'
[Wɔ́n dhɛ́].

Interestingly, *dhừn* is often used with nouns for abstract notions (names of actions, events, time spans, etc.), as in (30–32), despite a low position of these words in the Animacy Hierarchy.

(30) *Wɔ́n bhā yɛ́ kʌ̀ bhēn gbàn kɔ́n wɔ̀n ká, yɛ́*
   matter ART 3SG.JNT do\JNT human all astonish matter\IZF with CONS
*wó' **gbő-dhừn** bɔ̄ ...*
3PL.JNT\3SG.NSBJ crying-PL pass
'This matter amazed everybody, they began crying ...' [Wɔ́n dhɛ́].

(31) ***Yằkwʌ̀ʌ̀-dhừn** wà k̄ā, yáa dhó'-dhừn tằ sʌ̀ ká.*
   meeting-PL 3PL.PRF do 3SG.NEG.IPFV go\3SG.NSBJ-PL on good with
'When meetings were organized, he did not go there regularly' [Wáa wɔ́n].

(32) *Dhʌ́n sɛ́ɛndhʌ́n-dhừn zʌ̀ ằ-dhừn dhɔ̄ɔbhằa **wɔ̀n-dhừn** dhūn ɛ́*
   child small.PL-PL TOP 3SG.NSBJ-PL wish matter-PL FOC REL
wó bhā.
3PL.JNT there
'As for the children, there were some matters that amused them' [Wáa wɔ́n].

A special case is represented by names of body parts that form natural pairs (or natural sets, such as 'fingers' or 'finger-nails'). By default, they appear without the plural word, even if the complete set of body parts of the human being is meant, also if more than one person is concerned (33, 34). However, albeit more rarely, they may occur with the plural word (35). See also example (36), where among three occurrences of the same type, *dhừn* is absent in two instances and present in one.

(33) *Wɔ̀ yằ bhūn dɛ́dɛ̀, yɛ́ pʌ bhā'-dhừn v̀n yɛ̀*
   3PL.EXI sit\NEUT there a.little CONS thing ART\3SG.NSBJ-PL noise 3SG.EXI
*dằ ằ-dhừn **tő** gú.*
enter\NEUT 3SG.NSBJ-PL ear in
'They stayed there a little, and the noise of these things reached their ears'
[Wɔ́n dhɛ́].

(34) *Tῲ̰̋ŋ̀dā̰-zλ̀-dhὺn plὲ̀ wó wō kó sᷛη bhō-sīʌ,*
friend-PL.M-PL two 3PL.JNT 3PL.REFL RECP.CMM greeting remove-DUR
wὂ̀ wō kó kɔ̄ŋ-gā wὲ̀.
3PL.EXI 3PL.REFL RECP.CMM finger-bone speak\NEUT
'When two friends greet each other, they hook each other's fingers and snap'
[d165:18].

(35) *Ȁ̀ **kɔ̀-dhὺn** wὰ̀ gā' bhὰ.*
3SG.NSBJ hand.CMM-PL 3PL.PRF die\3SG.NSBJ on
'His hands do not function' (because of tiredness) [d163:21].

(36) *Yȁ̀ Gwēe **kwēᷛ᷄** dhὲ̄-gὰ̀, kɤ̄ yὰ̀' **gὲ̀n sūᷛ** gὰ̀, kɤ̄*
3SG.PRF Panther hand look SMLT 3PL.PRF\3SG.NSBJ foot claw look SMLT
*yὰ̀' kɔ̀ **sūᷛ-dhὺn** gὰ̀, yɤ̋' pɤ̄ …*
3PL.PRF\3SG.NSBJ hand.CMM claw-PL see CONS.3SG.JNT\3SG.NSBJ say
'He looked at Panther's front legs, he looked at his back leg claws, he looked
at his front leg claws, and he said …' [Wɔ́n dhɤ̋].

Analysis of the data available reveals a tendency (which needs to be further verified
on a larger corpus) reported earlier for another Mande language, Bambara (Vydrin
2016; Vydrin 2019: 314–315). When body parts fulfil their typical functions (ears hear
sounds; legs serve for walking or running; eyes see the surrounding world, etc.),
the name of the body part appears without the plural word. When they disfunction
or appear as passive objects, the plural word appears.

Declinable nouns (see §1.1) represent another special case. When used in the
common case, they are pluralized in the same way as inflexible nouns, by means
of the plural word *dhὺn* (37). However, these plurals rarely occur in natural texts.

(37) *Sᷛăkpᷛădā-dhὺn wὰ̀ zᷛ̈ **dhēᷓᷓ̄-dhὲ̄-dhὺn** ɤ̋ wó gūn*
mangoose-PL 3PL.PRF bee hive-CMM-PL REL 3PL.JNT be.PST\JNT
bᷛū gú yā ȁ̀-dhὺn gbὰ̀n bhɤ̀.
wilderness in here 3SG.NSBJ–PL entire eat
'Mongooses have ravaged all the beehives that had existed in this forest'
[f085:01].

In the oblique cases (resulting historically from the fusion of a nominal stem with
a postposition), the plural is not easily formed. If one insists, speakers of Dan do
produce some complex forms where oblique and common case markers are com-
bined, and the etymological postposition is "extracted" from the fused form. E.g.,
in the common case, the singular form of a declinable noun meaning 'edge (of a
field)' is *blᷛėedhὲ̀*, and its singular locative case form is *blᷛ̈edhɤ̄* (the elements -*dhὲ̀*
and -*dhɤ̄* being case suffixes). Its plural form in the locative case is *blᷛ̈edhɤ̄dhὲ̀-*

dhʋ̀n, this form (comprising both case suffixes) cannot be used any more without a postposition. Such forms look artificial and hardly ever occur in natural texts.

Some names of insects (those which usually appear in great numbers) have inherently collective meanings and are incompatible with the plural word *dhʋ̀n,* e.g. *zlőo* 'flying termites', *zlūu* 'driver ants'. However, one cannot extend this to all the names of insects (not even to all names of insects which live in collectives). Thus, *ző* may designate both a bee or a swarm of bees (a collective meaning), and *ző-dhʋ̀n* is used for the plural (38). *Zlʌ̄ʌndhʌ́n* 'small black ant(s)' may have a singular or a collective meaning, and *zlʌ̄ʌndhʌ́n-dhʋ̀n* is used for the plural (39).

(38) **Ző-dhʋ̀n** wà   dằ   ző   dhēɛdhɣ̄.
     bee-PL   3PL.PRF go.up bee hive.LOC
     'Honey bees entered the beehive' [f84:11].

(39) *Zlʌ̄ʌndhʌ́n-dhʋ̀n* wà   dɔ̄ sűkắdhű bhā'   bhằ.
     black.ants-PL   3PL.PRF put sugar   ART\3SG.NSBJ on
     'Black ants have clustered on the sugar' [e15:14].

2.3.6 Number deficiency in nouns: pluralia tantum and singularia tantum

Pluralia tantum are not attested in Eastern Dan. For pluralia tantum adjectives, see 3.1.2. Singularia tantum for the names of insects have been discussed in 2.3.5; for the other nouns, the identification of singularia tantum is difficult because of the facultative character of the plural word.

2.3.7 Mass/matter nouns

See 2.3.4 for a semantic interpretation of mass nouns with plural words.

2.3.8 Group and set nouns

There is no sufficient data.

2.3.9 Action nominals

In Eastern Dan, there are several kinds of action nominals: a gerund derived with the clitic *-sʋ̀*; a masdar derived with the suffix *-dhʋ̀*; a result name (of a limited productivity) derived with *-dhē*; a suffixless verbal noun. A verb can also be nominalized without any modification of its form in a construction with the dummy verbs *wō* or *bhō*.

Gerunds (40) and verbal nouns in constructions with dummy verbs (41) can be pluralized, although such occurrences are not very frequent in natural texts.

(40) *Yà*   *kā*   *kó*   ***zī̄-sṳ̀-dh̀ṳn*** *ká,*   *wằa*   *pɔ́n,*   *yɤ̀*
3SG.PRF do RECP beat-GER-PL with 3PL.PRF>3SG.NSBJ drop 3SG.SBJV
ɤ̄   *dè*   *tằ-kɔ̄ɔ̀,*   *yɤ̀*   *gó*   *ɤ̄*   *tɛ́́ɛdō* *wlɤ̀ɤ.*
REFL.SG self protect 3SG.SBJV leave REFL.SG friend under
'If there are wrestlings, and if he has been knocked down, he protects himself and gets out from under his partner' [Wáa wɔ́n].

(41) *Yà*   *lòo*   *pɤ̂*   *ő̋o* *pɤ̂*   *ɤ́'*   *gú,*   *ằ*   *t̀ʼŋ*   *dhʌn*
3SG.PRF arrive village any village REL\3SG.NSBJ in 3SG.NSBJ rumour FOC
dằ,   *yɤ́*   *dhʌndhīʌ́ŋ́-dh̀ṳn* *wò*   *kó*   ***zī̄-dh̀ṳn*** *wō* *ằ*   *tằ.*
go.up\JNT CONS girl-PL   3PL.SBJV RECP beat-PL do 3SG.NSBJ on
'When he came to any village, and when the rumour about him spread, girls fought over him' [Wɔ́n dhɤ́].

2.3.10 Nominalized adjectives

In Eastern Dan, adjectives cannot be used headlessly.

2.4 Verbal number

There are two formal means for expressing verbal number. Both are relatively infrequent in natural texts.

2.4.1 Full reduplication of the verb.

Most often, verbal number is expressed by full reduplication of the verbal stem. It may convey the meaning of participant number, i.e. an action involving more than one entity (42), or event number, i.e. multiple situations involving one entity (43); and both meanings can combine in one context (multiple situations involving multiple entities). Availability of the contexts like (43) proves that the full reduplication of verbs cannot be interpreted as an agreement with a plural argument.

(42) *Bhān*   *dhɤ̀n-dh̀ṳn* *wà*   *bhǐ̋a-bhǐ̋a* *wɛ̋́ɛwɛ̋́.*
1SG.POSS trap-PL   3PL.PRF snap-snap in.vain
'My traps have snapped in vain' [f129_09].

(43) *Bhɛ̄n   bhā yà   dlāā̀n-dlāā̀n.*
     human ART 3SG.PRF slide-slide
     'The man has slid (many times)' [f130_03].

The interpretation of the meaning of reduplication depends on both the lexical se-
mantics of the verb and the context. In some cases, reduplication can express mean-
ings other than verbal number, including long duration of the situation (44), its
intensity (45), distributivity or exhaustivity (46). Nevertheless, the pluractional
reading remains the default one for the fully reduplicated verbs.

(44) *Zân yà   bhɛ̄ɛ́-bhɛ̄ɛ́ Tɔ̀kpà dhɛ̀.*
     Zan 3SG.PRF beg-beg   Tokpa before
     'Zan was imploring Tokpa for a long time' [f129_07].

(45) *Bhá   bhān   pàŋ blǐ-blǐ.*
     2SG.PRF 1SG.POSS pants wear.out-wear.out
     'You have worn out my pants too much' [f128_08].

(46) *À̀   kwī yà   bhā-bhā.*
     3SG.NSBJ skin 3SG.PRF come.out-come.out
     'Rush has broken out over his entire body' [f129_04].

For those verbs that contain preverbs (see 2.4.2), the reduplication applies to the
entire stem, see (47) for the verb *yǎn-bhō* <eye-remove> 'solve'.

(47) *yɤ̀̀   wɔ́n   gbàn yǎn-bhɔ̀-yǎn-bhɔ̀.*
     3SG.EXI matter all   solve\NEUT-solve\NEUT
     'He solves all the matters (one by one)' [f146_05c].

Full reduplication is common, and, in elicitation, turns out to be available for a little
less than half of the verbal lexicon (with some semantic and lexical constraints).
     A reduplicated verb can appear in the same TAM constructions as a simple verb.
In (47), it occurs in a neutral aspect construction, and in the previous examples,
in the perfect construction. Therefore, the reduplication cannot be regarded as an
aspectual category.
     I have no information as to whether a reduplicated verbal form may refer to two
events. In any case, there is no special verbal marker for two events.
     A single (non-reduplicated) verbal form can also express a multiplicity of
events; it is semantically unmarked in this respect. There may however be a seman-
tic difference: a reduplicated verb with a plural argument tends to express a distrib-
utive meaning, and a single (non-reduplicated) verb is neutral in this respect.

2.4.2 Insertion of the plural word *dhʉ̀n* between the preverb and the verbal stem

Another means to express pluriactionality is limited to the verbs with preverbs. In Eastern Dan, preverbs are elements of nominal origin added to verbal roots on the left to derive new verbs. The preverbs are easily separable, and yet, there are strong arguments in favour of regarding a verb and its preverb as a single lexeme. One such argument is the possibility of a full reduplication of the verb including the preverb, see (46). For more details on preverbs in Eastern Dan, see (Vydrin 2009a), in Kla-Dan, (Makeeva 2013). For a discussion of the status of preverbs in a closely related Mano language, see (Khachaturyan 2017).

One of the elements that can be inserted between a preverb and its verbal stem is the plural word *dhʉ̀n*. Its insertion ("preverb pluralization" below) expresses a pluriactionality. Therefore, the scope of the pluralizer is the entire verb, see (48a, 48b) for the verb *dhíɤ̋-tó* <edge-leave> 'to interrupt'.

(48) a. *Yà yī̄ dhíɤ̋-tó.*
3SG.PRF work edge-leave
'He has interrupted a/the work'.

b. *Yà yī̄ dhíɤ̋-dhʉ̀n tó.*
3SG.PRF work edge-PL leave
'He has interrupted (different, several) works' [f150_05].

Just like the reduplication, preverb pluralization can refer to actions concerning many participants (48b) or to multiple events concerning one referent; see (49) for the verb *tȁ-kún* <on-catch> 'to help'.

(49) *Yà ɤ̄ bhā dhēbȁ tȁ-dhʉ̀n kún.*
3SG.PRF REFL.SG POSS woman on-PL catch
'He has helped his wife several times' [f149_09].

The preverb pluralization cannot be regarded as an agreement with an argument (cf. (49) where the argument is singular) and is not a part of the TAM system of verb (because the verbs with pluralized preverbs can appear in different TAM constructions).

A verb with a non-pluralized preverb is semantically unmarked with respect to pluriactionality, and can appear with plural arguments.

3 Agreement and the syntax of number

3.1 Adjectives

3.1.1 The attributive function: *dhɨ̀n* pluralization

By default, an NP containing an adjective[12] is pluralized by means of the plural word *dhɨ̀n,* which appears to the right of the adjective (50a). In an attributive construction (where an adjective follows the modified noun), *dhɨ̀n* can also appear after the noun (50b) or both after the noun and after the adjective (50c); the latter case could be interpreted as agreement if *dhɨ̀n* were a suffix and not a plural word. There seems to be no semantic or pragmatic difference between these variants; however, the first strategy (50a) is by far the most frequent one, and the other two (50b, 50c) are less current.

(50) a. *kɔ́   zɔ̈̀œndhē-dhɨ̀n*
     house red-PL

   b. *kɔ́-dhɨ̀n   zɔ̈̀œndhē*
     house-PL red

   c. *kɔ́-dhɨ̀n   zɔ̈̀œndhē-dhɨ̀n*
     house-PL red-PL
     'red houses'

3.1.2 Pluralization through modification of the stem

The plurality of an NP can be expressed through modifications of the adjective stems, which is less straightforward and characterized by the following irregularities:

1. The modification of stems manifests itself in reduplication, tonal changes and/ or transfixation. The modification pattern depends on the structure of the initial form of the adjective. Adjectives can be grouped into several formal classes ("phonotypes"), but even within classes, considerable diversity of patterns is observed.
2. Modifications of adjectival stems express plurality cumulatively with the category of intensity. Adjectives in Eastern Dan can have up to three degrees of intensity, and each of these meanings can be expressed together with the plurality. Each adjectival phonotype has a different paradigm of intensity-plurality.

---

12 Morphology and syntax of adjectives in Eastern Dan are analyzed in detail in (Vydrine 2007; Vydrin 2009b). In this section, only a short survey of the plural formation for adjectives is provided.

I establish five main "phonotypes" and several intermediate cases. Tables 6–9 show some examples of the plural-intensive paradigms (empty cells stand for non-existent forms); for more details, see (Vydrine 2007; Vydrin 2009c).

Tab. 6: Phonotype 2 (CVV with different tones).

Basic form	Singular intensive	Plural intensive	Singular superintensive
kpíî̀ 'big'	kpíkpì̀		kpíkíkpì̀kì̀
gbínɲ̀ 'heavy'	gbínŋgbìnŋ	gbínɲ̀gbínɲ̀	gbíŋkíŋgbìnkìn
klɤ̀ɤ̌ 'short'	klɤ̀ɤ̌	kpɤ̌klɤ̌ ~ klɤ̌ɤ̌klɤ̌ɤ̌	

Tab. 7: Phonotype 3 (CVV with two identical tones).

Basic form	Plural intensive	Singular superintensive	Plural superintensive
tīī 'black'	tītī	tīitïdhè̀	tīitīī
zīī 'old'	zīzī	zīizïdhè̀	zīizīī

Tab. 8: Phonotype 5 (CVVCVV, tones eH-eH-eL-eL).

Basic form	Singular intensive	Plural intensive	Singular superintensive
dhɛ́ɛ́ndhɛ̀ɛn 'tasty'		dhɛ́ɛ́ndhɛ́ɛn	dhɛ́nḱɛndhɛ̀nkɛ̀n
plʌ́ʌ́plʌ̀ʌ̀ 'light', 'soft', 'tasty'		plʌ́ʌplʌ́ʌ 'very tasty (pl.)'	plʌ́ḱʌplʌ̀kʌ̀
slʌ́ʌnslʌ̀ʌn 'sweet'	slʌ́nkʌ́nslʌ̀nkʌ̀n	slʌ́ʌnslʌ́ʌn	

An adjective sɛ̄ɛ́ndhʌ́n 'small' represents an individual type; its basic plural form is derived from the basic singular form by a tonal modification alone.

Tab. 9: Plural/intensive derivation of the adjective sɛ̄ɛ́ndhʌ́n 'small'.

	Plural	Singular intensive	Plural intensive	Singular superintensive
sɛ̄ɛ́ndhʌ́n 'small'	sɛ́ɛndhʌ́n	sɛ̄ɛ́ɛ̀ndhʌ́n	sɛ́nsɛ́ndhʌ́n	sɛ̀ɛsɛ̄dhè̀

Most of plural forms derived through modifications of the stems can be optionally accompanied by the plural marker dhʉ̀n.

There are two suppletive plural forms: sʌ̀ 'good' – plural intensive sɛ̀ŋbʌ̀, plural extraintensive bhɛ̆bhɛ̆.

Two adjectives are pluralia tantum: glɔ̆ɔglɔ̆ɔ 'robust and tall.PL (of young people)' and gblɤ̆gblɤ̆ 'large.PL'. Both can appear with or without the plural word dhʉ̀n.

Here are the reasons to believe that in both cases we have pluralia tantum, rather than suppletive plural forms:

Glɔ̋ɔglɔ̋ɔ with humans means 'robust, tall and young.PL', and with names of in-animate objects, 'medium-sized.PL'. There is no adjective in Eastern Dan (to my knowledge) for the complex meaning 'robust, tall and young' in the singular; i.e., the meanings included in the lexical meaning of *glɔ̋ɔglɔ̋ɔ* can be expressed (some-times with nuances) in singular, but they are accumulated in one lexeme only in the plural. The notion 'medium-sized' in the singular seems not to be lexicalized either.[13]

Gblɏ̋gblɏ̋ 'large, big.PL' is used with or without the plural word *dhṳ̀n*. On semantic grounds, it can be related to *kpîì* 'big, large'. *Kpîì* has a regular plural form, *kpîì-dhṳ̀n,* which makes interpretation of *gblɏ̋gblɏ̋* as a suppletive plural form of *kpîì* con-troversial (although not totally excluded). In texts, *gblɏ̋gblɏ̋* is used much more fre-quently than *kpîì-dhṳ̀n.*

3.1.3 Pluralization of adjectives in predicative function

In Eastern Dan, there are two constructions for the predicative use of adjectives, one involving the comitative postposition *ká* (51), the other without it (52).[14]

(51) *yɤ̏   blɏ̋ŋblɏ̏ŋ ká   dhɛ̀ dhṳ̀aŋdhɛ̏ dhɤ́.*
   3SG.EXI lustrous with that mirror   be
   'It shines (lit.: it is shining, it is lustrous) like a mirror' [h09_03].

(52) *Kpìʌʌ   yɤ̏   yâ.*
   chimpanzee 3SG.EXI bad
   'A chimpanzee is ugly' [c029_06].

I have systematically checked all adjectives attested in Eastern Dan for their number agreement in the predicative function. It turns out that there is a crucial difference between the two constructions in this respect. In the construction without the post-position, an adjective usually cannot agree in number (53), while in the construction involving the postposition *ká,* the agreement is usually allowed or even obligatory.[15]

---

13 There is a derived adjective *zìnŋgúsṳ̀* 'middle' (e.g., *gēnŋgā zìnŋgúsṳ̀* 'middle toe'), but it des-ignates a position, and not a size.
14 I present here the predicative use of adjectives only in non-verbal clauses. They can also appear in depictive verbal constructions (as secondary predicates), both with the postposition *ká* or with-out it. Unfortunately, I do not have enough data on the use of different adjectives in depictive constructions (in particular, with respect to their pluralization).
15 My language assistant accepted pluralization for only two adjectives in the construction without *ká*: *gbɛ̋kɛ́gbɛ̀kɛ̀sṳ̀-dhṳ̀n* 'extremely difficult' (a focalized superintensive form derived from *gbɛ̏ɛ̀* 'dif-ficult') and *glɛ̀ŋsṳ̀-dhṳ̀n* 'smart'. He also approved pluralization of about 4/5 of all the adjectives in the construction with *ká.*

The plural can be expressed by the plural word *dhɨ̀n* (54) or by modification of the stem, in this case an adjective can appear even in the postpositionless construction (55).

(53) *Īī*   *kɨ̄-kɔ̀-dhʌ́n-dhɨ̀n*   *wɔ̀*   *kʌ́nŋkʌ́nŋ/kʌ́nŋkʌ́nŋ-dhʉ́n.*
     2SG.NSBJ do-hand.CMM-DIM-PL 3PL.EXI bumpy/*bumpy-PL
     'Your little ways of doing are doubtful' [c035_14].

(54) *À*   *bhǎ*   *dhēbʌ̀-dhɨ̀n* *wɔ̀*   *bhʌ́ŋ̂bhʌ̀ŋ-dhɨ̀n* *ká.*
     3SG.NSBJ POSS woman-PL   3SG.EXI plump-PL   with
     'His wives are plump'.

(55) *Gbên dhɛ́-dhɨ̀n*   *wó*   *pʉ́pʉ́*   *dhʌ̀n wà*   *dhūn.*
     dog   the.other-PL 3PL.JNT white.INT.PL FOC   3PL.PRF come
     'These are dogs who are white who came' [d189:3].

3.2 Number agreement in predicative markers

As mentioned in §1.1, in Eastern Dan the subject is indexed on the Aux (auxiliary words, syntactic heads of clauses, in the Mandeist tradition referred to as "predicative markers"; some of which can also appear in non-verbal clauses and assume copula functions). Such indexation can be viewed as person/number agreement.

     The predicative marker is indexed as plural if the coreferent NP is formally pluralized by the plural word *dhɨ̀n* (56).

(56) ... *yɤ́*   **dʉ̀ɤ-dhɨ̀n wó**   *dhō*   *wó*   *tlőo*   *bhā'*
      CONS buffalo-PL 3PL.JNT go\JNT 3PL.JNT termite.hill ART\3SG.NSBJ
     wūu.
     break\JNT
     '... and the buffaloes went and broke the anthill' [Wɔ́n dhɤ́].

If the head of a subject NP has no plural marking, the predicative marker is usually singular, even if it refers to a plural object, as in (57), where legs of numerous animals are meant, but the word *gɛ̀n* 'foot, leg' is not accompanied by the plural word *dhɨ̀n*.

(57) *Wū*   *dhɛ́-dhɨ̀n kɤ́*   *ʌ̀-dhɨ̀n*   *gɛ̀n yáa*   *gbɛ̋ɛ̀* ...
     animal other-PL   REL 3SG.NSBJ-PL leg 3SG.NEG.IPFV hard
     'Some animals whose legs are not very strong ...' [Wáa wɔ́n].

However, disagreement in number is often observed when the subject NP is relativized. In this case, it may have no plural marker, while the predicative marker is plural.

(58) **Bhɛ́n** **wó** dằ kɔ́ɔdhɤ̄, wāằ **bhɛ́n** **wó**
   human\REL 3PL.JNT go.up\JNT house.LOC 3SG.and human\REL 3PL.JNT
   tō kpɛ́nŋdhɤ̄, ằ-dhừn dhɯ́ dō.
   remain\JNT outside.LOC 3SG.NSBJ-PL number one
   'The number of those who have entered the house and of those who have remained outside is the same' [g19_08].

Plurality can also be expressed by the quantifiers *gbɛ̂* 'many' (59), *tɛ̀tɛ̀* 'many', *gbàn* 'all' (60), *pɛ́pɛ́* 'all, each', *dódó* 'few', *kɛ́ɛ* 'few' (61), *tɤ́ŋbhằsừ* 'numerous' (in a negative context), *kpí̂dhɛ̀* 'majority'. An NP with a quantifier requires a plural predicative marker.

(59) *Bhēn gbɛ̂ wà lòo kó tằ fɛ̂tɤ̀ kā-yằn gú.*
   human numerous 3PL.PRF arrive RECP.CMM on holiday do-SPN in
   'Many people have gathered for the celebration' [f60:8].

(60) *Bhēn gbàn wó dhɤ́ wà dhūn dhɛ́ɛdhɤ̄.*
   human entire 3PL.JNT be 3PL.PRF come village.square.LOC
   'Everybody (lit.: all the people who exist) has come to the public square' [f185:14].

(61) *Bhēn kɛ́ɛ wà dhūn.*
   human few 3PL.PRF come
   'Few people have come' [g41_01].

3.3 The use of plural word inside an NP with a quantifier

Most often, an NP followed by a quantifier has no plural word (59–61). However, here again, it is not a rule but a tendency. Counter-examples, however infrequent, can be found (62, 63). The semantic or pragmatic contribution of the plural word in this context is not yet clear.

(62) **Gɔ̄n-dhừn gbɛ̂** **wɔ̀** sìɛtíī̂ bhừn wó
   male-PL numerous 3PL.EXI cigarette drink\NEUT 3PL.JNT
   zɩ̀ɤ' ká dhēbằ-dhừn tằ.
   pass\JNT\3SG.NSBJ with woman-PL on
   'More men than women smoke' [qk157].

(63) **Bhīnń-dhừn gbàn** *dhūn* **wáa** *súǹ* *bhán.*
Jula-PL entire FOC 3PL.NEG.IPFV Ramadan put
'Not all Jula[16] observe Ramadan' [qk129].

There is also an idiomatic context where the NP is necessarily followed by the plural word and by the quantifier *gbàn* 'all'. If this NP occupies the subject position, the predicative marker indexes plurality. The construction in question conveys the meaning of a mismatch between the pragmatic expectations and the reality (an approximate equivalent of the English *even*), as in (64, 65), or the meaning 'and even more …', 'and on top of everything', as in (66).

(64) *Gó ǹ gɔ̏ á zīɤ, wɔ́n á dɔ̄' ká,*
leave 1SG.NSBJ PP 1SG.SBJV pass matter 1SG.JNT put\JNT\3SG.NSBJ with
*ɤ́ **bīɤ-dhừn** **gbàn** wà fìʌʌ' bhȁ yā, bhī áa*
when elephant-PL all 3PL.PRF fail\3SG.NSBJ on here 2SG.AUT NEG
bhūn í dhō slòo' bhȁ.
there 2SG.JNT go\JNT suffice\INF\3SG.NSBJ on
'Let me go, the matter that worries me, even the elephant has failed to settle it, you won't solve it' [Wɔ́n dhɤ́].

(65) *yɤ̏ sɔ̋ dhȉɤ **lɔ̄ɔ-dhừn** **gbàn** ká.*
3SG.EXI trap set\NEUT grass.sp-PL all with
'He makes traps even of the "loo" grass' [g42_13].

(66) **Kplǎtő-dhừn gbàn** *wȁ bhȁn à gɔ̏.*
bald.patch-PL all 3PL.PRF beat 3SG.NSBJ PP
'And on top of everything, he has grown bald' [f185_05] (lit.: 'All the bald patches have struck him on the head').

3.4 Patterns of agreement with numeral modification

As with all other modifiers, numerals (being determiners) follow the NP. By default, the modified NP appears in its bare form (i.e. without the plural word *dhừn*).

(67) *Yà sɔ̋ɔ yǎagā zī.*
3SG.PRF cane.rat three kill
'He has killed three cane rats' [d085_14].

However, constructions with numerals can contain the plural word *dhừn*. These constructions are associated with specific meanings, as explained below.

---

16 Jula is a name of an ethnic group in Côte d'Ivoire.

3.4.1 Plural word after an NP with a numeral

The plural word can be inserted between a noun and a numeral to express definiteness.

(68) **Wɔ̄kɔ̏-dhṳ̀n kɔ̀ŋ dō ɤ̄ gā plɛ̀** *bhā, ǎ-dhṳ̀n klɔ̏ɔ̏ dhɛ̀n.*
rule-PL ten one REFL.SG bone two ART 3SG.NSBJ meaning here.is
'Here is the meaning of the twelve rules' [Wáa wɔ́n].

(69) **Dhī́dhǎan bhán bhɛ̀n-dhṳ̀n plɛ̀** *bhā wà wō kó pɤ̀*
quarrel put human\IZF-PL two ART 3PL.PRF 3PL.REFL RECP.CMM fall
wɔ̀ slʌ̄ʌ-sīʌ wō gú bɔ̀bɪ̀ gú.
3PL.EXI turn-DUR PL.REFL in dust in
'Both quarellers have brought each other down and are rolling in the dust'
[f082:06].

3.4.2 Plural word after a numeral: exactness and approximation

The plural word can follow the numeral. This may express two different meanings:
the exactness of the quantity expressed by the numeral (70), or its opposite, approximation (71);[17] my data are insufficient to formulate a rule of distribution for these two readings.

(70) *Pōlísɤ̀-dhṳ̀n wà* **kwànbhɛ̀n yǎagā-dhṳ̀n** *kún.*
policeman-PL 3PL.PRF thief three-PL catch
'The policemen have caught exactly three thieves' [f181_02] (the context: we don't know if these thieves are the same who had evaded earlier, but their number is the same).

(71) *... yɤ́ zūn kpɛ́ŋ bɛ̀ŋ ká, kɤ̄ bhlǎntó kpɤ̄ dō*
3SG.JNT reach\JNT outside.IN running with SMLT hammer ball one
ɤ́ bhɔ̀ɔn **kɪ́dhő́ŋ sɔ̏ɔdhű-dhṳ̀n** *bhǎ yǎ gɔ̏.*
REL.3SG.JNT be.able\JNT kilo five-PL on 3SG.EXI\3SG.NSBJ PP
'... he was running outside with a hammer of about five kilos in weight'.

---

17 Cf. (Corbett 2000: 230–240) on the approximative meaning of the plural in some languages. Note that in languages quoted there, it is a noun that is marked for plural. In Eastern Dan, it is the numeral.

3.4.3 Numeral modification and number agreement in predicative markers

When pluralization of a subject NP is conveyed by a numeral, its agreement in number with the predicative marker is determined by the Animacy Hierarchy. With human and animate subjects, a plural predicative marker is preferable, although a singular predicative marker is possible (72). With inanimate and abstract subjects, singular predicative markers predominate (73), although agreement in number can be found even for such nouns (74).

(72) *Bhēn   sɔ̋ɔdhű̋ **wȍ ~ yȑ***   *sʌʌdhɛ̋ dlāǎn-sȕu bhǎ kwɪ́plȑɤ.*
     human five   3PL.EXI ~ 3SG.EXI paper teach-GER POSS town.LOC
     'Five people do their studies in the city' [f33_03].

(73) *Dhű̋u kwēɛ̋ plɛ̀ **yȑ**   gìi-sɪ̄ʌ   kwʌ̋ʌ̀   víȍovǐooǐoodhɤ̄.*
     tree hand two 3SG.EXI wound-DUR RECP.SUP creaking
     'Two branches of a tree rub against each other with creaking' [e29:11].

(74) *Kʌ̋flɛ̋e dhɛ̋ yǎagā **wǎ**   pȑ.*
     coffee leaf three 3PL.PRF fall
     'Three leaves of the coffee tree have fallen' [g13_03].

I have no information concerning coordination of numerals.
    Bare numerals can act as predicates (75, 76).

(75) *Dhē   yȑ   gùn'   gɔ̏ **kȍ̀œŋ plɛ̀,** kɛ̋ɛ yɤ́   dhēbɔ̏*
     woman 3SG.EXI be.PST\3SG.NSBJ PP ten   two but CONS woman
     bhā'   bhá yíi   dhʌ́n kpɔ́.
     ART\3SG.NSBJ certain 3SG.NEG.PST child spread
     'He had twenty wives (lit.: 'woman was with him twenty...'), but none
     of the wives gave birth to a child' [Wɔ́n dhɤ́].

(76) *Yȑ   dhó klàŋ   dhɪ́ɤ kɤ̄'   kwɛ̏ **yȑ** sɛ̋œ̀̀plɛ̀.*
     3SG.EXI go   formation before SMLT\3SG.NSBJ year 3SG.EXI seven
     'He went to school when he was seven years old' (lit.: 'He went to
     the school, and his year is seven') [e20:13].

3.5 Coordination

On the expression of coordination in Eastern Dan, see §2.2.3. If conjoined NPs appear in the subject position, the predicative marker is necessarily plural, i.e. the

entire coordinative construction is the controller of the agreement (77). This is true even if the conjoined NPs are represented by inanimate nouns (78).

(77) *T̀ òkp̀ à wāa Zân ẁ à Yɔ̀ tằkó-zī̄.*
NOM.M 3ICL NOM.M 3PL.PRF NOM.F contest
'Tokpa and Zan contest Yo' [h10_14].

(78) *Dhű̀u wāằ kɔ́ ẁ à pɤ̀.*
tree 3ICL house 3PL.PRF fall
'A tree and a house fell down' [f148:13].

3.6 Number in anaphoric relations

By default, plural pronouns are used when their antecedent NPs are marked for plural (79). However, mismatches are also frequent. An NP without the plural word can be referred back to by a plural pronoun; cf. an example of such semantic agreement in (80). In the same way, a plural predicative marker may correspond to a formally non-pluralized NP located in another clause: in this case, a *pro* can be postulated (81).

(79) … *yɤ́ kpằn **dừ ɤ-dhừn**ᵢ bhā'-**dhừn**ᵢ kpɤ̄ bhằ.*
CONS.3SG.JNT see\JNT buffalo-PL ART\3SG.NSBJ-PL crowd on
'… and he saw a group of buffaloes' [Wɔ́n dhɤ́].

(80) ***Bhɛ́n**ᵢ ɤ́ bhēn yáa'-**dhừn**ᵢ gɔ̀, ẁ à tó*
human\REL REL human 3SG.NEG.IPFV\3SG.NSBJ-PL PP 3PL.PRF remain
wʌ̀ʌsừ.
pitiful
'Those who have no such relatives remain unhappy' [Pamε1_1klaŋ_bha.005].

(81) [***Bhɛ́n**ᵢ ɤ́ dhɤ́*] [pro ᵢ *ẁ à dhūn dhɛ̋ɛdhɤ̀*].
human\REL REL\3SG.JNT be 3PL.PRF come village.square.LOC
'Everybody has gathered on the public square' [f185_16] (lit.: "Human who exists, they came…").

4 Semantics and discourse

Predicative markers indexed for third person plural are regularly used in impersonal sentences.

(82) *Wɔ̀   gbɔ̄ kʌ̀   gbɛ̄ɛ́   ká.*
3PL.EXI pot do\NEUT potter.clay with
'One makes pots from clay' [g03_19], lit. "they make a pot …".

(83) *Wɑ̀   yʌ̄ gbàn dhūn ū̄   dhɛ̏,   ɑ̏   kʌ̄.*
3PL.PRF work all   give 2SG.NSBJ before 3SG.NSBJ do
'Do all jobs that are offered to you' [g18_10], lit. "All the jobs they have
given to you, do it".

In generic sentences, a noun is usually used in the singular (unmarked) form, i.e.,
without plural word *dhʌ̀n* (84, 85); however, a plural form is also possible (86).

(84) *Gā-bhɛ̏n   áán   dhūn.*
death-human\IZF 3SG.GN.NEG приходить
'Dead people do not come back' or 'No dead person comes back'.

(85) *Yɤ̏   **kɔ́**   g̏ìi   sʌ̀.*
3SG.EXI house coat\NEUT well
'He coats houses well' [e35_04c].

(86) *ɤ́   gūn   kpá   pɑ̆pɔ̆ ɛ̆ɛn yɛ̆   dhūn ɤ́   wó*
3SG.JNT be.PST\JNT formerly panel or thatching.grass FOC REL 3PL.JNT
*gūn   **kɔ́-dhʌ̀n** tɑ̏-kʌ̄-sīʌ'   ká.*
be.PST\JNT house-PL thatch-DUR\3SG.NSBJ with
'In olden times, houses were roofed with "papot" panels or with straw'
[c032_05].

5 Conclusions

Number marking in Eastern Dan may appear simple and straightforward. As I tried
to show in this paper, this impression is misleading.

Among the irregularities, many are in line with the trends known from the pre-
vious studies of number, including some explained by the Animacy Hierarchy as
applied in Corbett (2000).
– At the top level of the Hierarchy, 1st pers. pronouns distinguish three number
   values, singular, dual and plural.[18] Another split of the pronominal system is
   between 1st & 2nd persons and 3rd person: the pronouns for interlocutors show

---

18 Alternatively, 1PL.DUAL can be interpreted as a singular inclusive pronoun; both these interpre-
tations make typological sense, cf. § 2.2.1.

number suppletion, while in the 3rd person, the number is marked by the plural word *dhʉ̀n* also used with nouns. This split is attested only in the non-subject series; in the other series, all plural pronouns are suppletive.

– Again at the top of the Animacy Hierarchy, but lower than the personal pronouns, we find certain irregularities in the plural formation of agent nouns and some kinship terms, see § 2.3.2.
– Plural marking is relatively common with the names of humans. In the lower parts of the Animacy Hierarchy, the plural word is less often used, and its absence does not necessarily mean singular. In Corbett's terms, the unmarked form can be interpreted as general number.

There is a correlation between case and number marking: in oblique cases, plural marking on the declinable nouns, the only class of nouns that distinguishes cases, is difficult or impossible.

Multiple marking is typical of some human names (including a couple of kinship terms, see § 2.3.2) and of adjectives, where plurality can be marked by a combination of the standard plural word *dhʉ̀n* and a modification of the stem. The latter expresses, together with number, degrees of the intensity of quality.

There are two mechanisms to express verbal number. One of them, the full reduplication of the verb, is typologically common, while the other one, pluralization of the preverb, is unusual. Neither of these mechanisms can be regarded as agreement with arguments; these are cases of true pluractionality.

Acknowledgments

Abbreviations and glosses

ART	definite article *bhā*
Atr	attribute
AUT	autonomous pronominal series
AUX	auxiliary
CMM	common case
CONS	consecutive conjunction (*yɤ́*)
DIM	diminutive suffix
DU	dual
DUR	durative verbal suffix *-sīʌ*
EMPT	empathy marker (*ʋ̀o*)

EXI	existential series of PM
EXCL	1 pers. exclusive plural pronoun/PM
FOC	a) focalization particle (*dhλn ~ dhūn*); b) grammatical high tone on certain focalized nouns
GER	gerundive suffix *-sʉ̀*
GN	gnomic series of predicative markers
ICL	inclusory pronoun
IMP	imperative series of PM
Incl.	inclusive pronoun
IN	inessive case
INF	infinitive marker (extra-low tone suffixed to the verb stem)
INT	intensive
IPFV	imperfective
IZF	isaphet marker (an extra-low tone on the noun)
JNT	conjoint series of PM; tonal modification on the verbal stem in the conjoint construction
LOC	locative case
MSD	suffix of masdar (verbal noun, *-dhɛ̏*)
N	noun
NEG	negative
NEUT	neutral aspect marker (extra-low tone on the verbal stem)
NMLZ	nominalization marker (extra-low tone)
NOM.F	female proper noun
NOM.M	male proper noun
NOM.MF	male or female proper noun
NP	noun phrase
NSBJ	non-subject pronominal series
O	direct object
PFV	perfective
PL	plural
POSS	possessive marker (*bhʌ̀, gɔ̋*); possessive pronoun
PP	postposition with a broad meaning (*gɔ̋*)
PM	pronominal marker (auxiliary)
PRF	perfect series of PM
PROH	prohibitive series of PM
PROS	prospective series of PM
PST	past
RECP	reciprocal pronoun
REFL	reflexive pronoun
REL	relativization marker (*ŷ*)
S	subject
SBJV	subjunctive series of PM
SG	singular
SMLT	a multifunctional conjunction (*kŷ*)
SPN	supine marker (*yǎn*)
SUP	superessive case or postposition (*bhʌ̀*)
TAM	tense – aspect – mood
V	verb
X	post-verbal arguments and adjuncts

References

Aplonova, Ekaterina. 2020. Dictionnaire goo – français. *Mandenkan* 63. 3–67.

Beam de Azcona, Rosemary. 2006. A Zapotec contribution towards the typology of inclusory constructions. Paper presented at the Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas. University of New Nexico. 5–8 January 2006.

Bearth, Thomas. 1971. *L'énoncé toura*. Norman (Oklahoma): Société Internationale de linguistique.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Daniel, Michael & Edith A. Moravcsik. 2013. The Associative Plural. (Ed.) Matthew S. Dryer & Martin Haspelmath. *The World Atlas of Language Structures Online*. http://wals.info/chapter/36 (14 September, 2020).

Erman, Anna. 2005. Le grammaticalisateur -ga en dan-blo. *Mandenkan* 41. 41–61.

Erman, Anna & Japhet Kahoué Loh. 2008. *Dictionnaire Dan–Français (dan de l'Ouest) avec un index français-dan*. St Pétersbourg: Nestor-Istoria.

Greenberg, Joseph H. 1988. The first person inclusive dual as an ambiguous category. *Studies in Language* 12(1). 1–18. https://doi.org/10.1075/sl.l2.1.02gre.

Khachaturyan, Maria. 2017. Bound noun plus verb combinations in Mano. *Mandenkan* 57. 3–24.

Khachaturyan, Maria. 2019. Inclusory pronouns in Mande: The emergence of a typological rarum. *Folia Linguistica* 51(3). 87–123. https://doi.org/doi.org/10.1515/flin-2019-2005.

Makeeva, Nadezhda. 2013. Les préverbes en kla-dan. *Mandenkan* 50. 85–102. https://doi.org/doi:10.4000/mandenkan.254.

Makeeva, Nadezhda. 2017. Kla-dan jazyk (Кла-дан язык) [Kla-Dan]. In Valentin Vydrin, Yulia Mazurova, Andrej Kibrik & Elena Markus (eds.), *Jazyki mira: Jazyki mande (Языки мира: Языки манде) [Languages of the world: Mande languages]*, 617–679. St. Petersburg: Nestor-Historia.

Mauri, Caterina & Andrea Sansò. 2018. Linguistic strategies for ad hoc categorization: theoretical assessment and cross-linguistic variation. *Folia Linguistica Historica* 39. 1–35. https://doi.org/10.1515/flih-2018-0001.

Vydrin, Valentin. 2009a. Preverby v jazyke dan-gweetaa (Превербы в языке дан-гуэта) [Preverbs in Dan-Gwèètaa]. *Voporsy jazykoznanija* 2. 75–84.

Vydrin, Valentin. 2009b. Morfologija prilagatel'nykh v dan-gweetaa (juzhnyje mande) (Морфология прилагательных в дан-гуэта (южные манде)) [Morphology of adjectives in Dan-Gwèètaa (South Mande)]. *Vestnik Sankt-Peterburgskogo universiteta* (Oriental and African Studies) 2. 120–140.

Vydrin, Valentin. 2009c. Morfologija prilagatel'nykh v dan-gweetaa (juzhnyje mande) (Морфология прилагательных в дан-гуэта (южные манде)) [Morphologie des adjectifs en dan-goueta (mandé-sud)]. *Vestnik Sankt-Peterburgskogo universiteta* (Etudes orientales et africanistes) 2. 120–140.

Vydrin, Valentin. 2010. Co-ordinative pronouns in Southern and South-Western Mande: A second compound pronouns area in Africa? In Konstantin Pozdniakov, Valentin Vydrin & Alexander Zheltov (eds.), *Personal pronouns in Niger-Congo languages: International workshop. St. Petersburg, September 13–15, 2010. Abstracts and papers: Working materials*, 164–172. St. Petersburg: St. Petersburg University Press.

Vydrin, Valentin. 2011. Déclinaison nominale en dan-gwèètaa (groupe mandé-sud, Côte-d'Ivoire). *Faits de langues: Les Cahiers* 3. 233–258.

Vydrin, Valentin. 2016. "On podn'als'a na nogu" vs. "jemu nadrali ushi": upotreblenie singul'arnykh i pl'ural'nykh form nazvanij parnykh chastej tela v bamana ("Он поднялся на ногу" vs. "ему надрали уши": употребление сингулярных и плюральных форм названий парных частей тела в бамана) [Use of singular and plural forms for the names of paired bodyparts in Bamana]. In Victor Vinogradov, Antonina Koval, Maria Kosogorova & Andrey

Shluinsky (eds.), *Issledovanija po jazykam Afriki (Исследования по языкам Африки) [Studies in languages of Africa]*, 22–35. Moscou: Kliutch-S.

Vydrin, Valentin. 2019. *Cours de grammaire bambara*. Paris: Presses de l'INALCO.

Vydrine, Valentin. 2007. Les adjectifs en dan-gwèètaa. *Mandenkan* 43. 77–103.

Angelika Jakobi and Gerrit J. Dimmendaal

3 Number marking in Karko and Nilo-Saharan

Abstract: Number distinctions in Karko are attested in all major lexical categories. Number values attested on nouns include singular, plural, transnumeral, singulative, and plurative, the selection of a specific number marking suffix often being determined by semantically defined groups of nouns. Adjectives, demonstratives, participles, and pronouns show a contrast between singular and plural (or nonsingular); personal pronouns additionally express clusivity (associated with the 1st person plural).

Verbal number, as realized by singular and plural stems, plays an important role in the morphosyntactic system of Karko; these stems interact with event number and with the number of the S or P participant. They also instantiate grammatical relations in ditransitive clauses, where the selection of a singular or plural stem is sensitive to the number of the direct object or the indirect object and their respective semantic roles. Additionally, plural stems are required to encode transitivity alternations. In derived intransitive clauses, being associated with low transitivity, plural verb stems convey passive, facilitative, and reflexive readings.

In terms of number marking Karko tends to be typical of Nilo-Saharan. Where additional phenomena occur for specific categories elsewhere in this phylum, this is illustrated in the article.

1 Overview

Karko (kko, glottocode 1256, ISO 639–3) is a Nubian language spoken in six small villages in the northwestern Nuba Mountains of southern Kordofan, Sudan. In her sociolinguistic survey, Krell (2012: 11) estimates that the population of the Karko villages amounts to 8,000 people; however, not all of these speak Karko. The *Ethnologue* (Lewis, Simons, and Fennig 2015: 244) gives the figure of 7,000 people speaking Karko as their first language. Observing that children speak mainly Arabic rather than Karko, Krell (2012: 11) concludes that speakers are shifting to Sudanese Arabic, which is the lingua franca of Sudan. The *Ethnologue* also assigns to Karko the status of a shifting language. The fact that Karko is not being fully transmitted to the next generation suggests that the language is highly endangered. This is equally true for many other languages in the Nuba Mountains. Therefore, Map 2 reflects the situation some 70 years ago, when Roland Stevenson (1915–1991) carried out linguistic surveys in the Nuba Mountains.

The Nuba Mountains are renowned for their linguistic heterogeneity; the languages spoken there belong to various branches of Nilo-Saharan, Niger-Congo and Kadu. Nubian constitutes a well-defined subgroup within the Eastern Sudanic
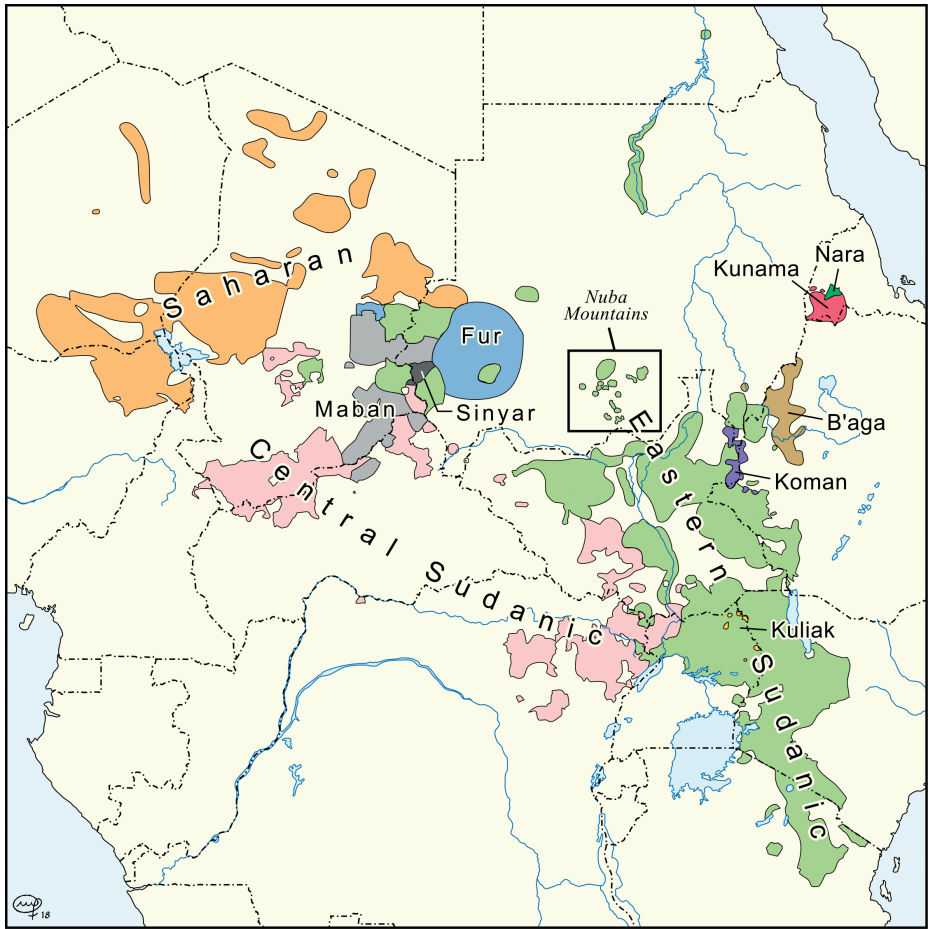
Map 1: The areal distribution of Nilo-Saharan.

branch of the Nilo-Saharan phylum (see Figure 1). Eastern Sudanic constitutes the largest subgroup within Northeastern Nilo-Saharan, which further includes Maba, Fur and Amdang, Kunama, Saharan, and Kuliak (Dimmendaal et al. 2019); Northeastern Nilo-Saharan is one of the two primary branches of Nilo-Saharan, the other primary branch being Central Sudanic.[1] The areal distribution of Nilo-Saharan is shown in Map 1.

In terms of its number-marking system, Karko tends to be a typical representative of Northeastern Nilo-Saharan. Where it differs from more general patterns in the phylum or subgroup to which it belongs, Eastern Sudanic, or where some other

---

1 Disputed members of Nilo-Saharan, Songhay, and the Koman or B'aga group are not considered here; see Dimmendaal et al. (2019) for further details.
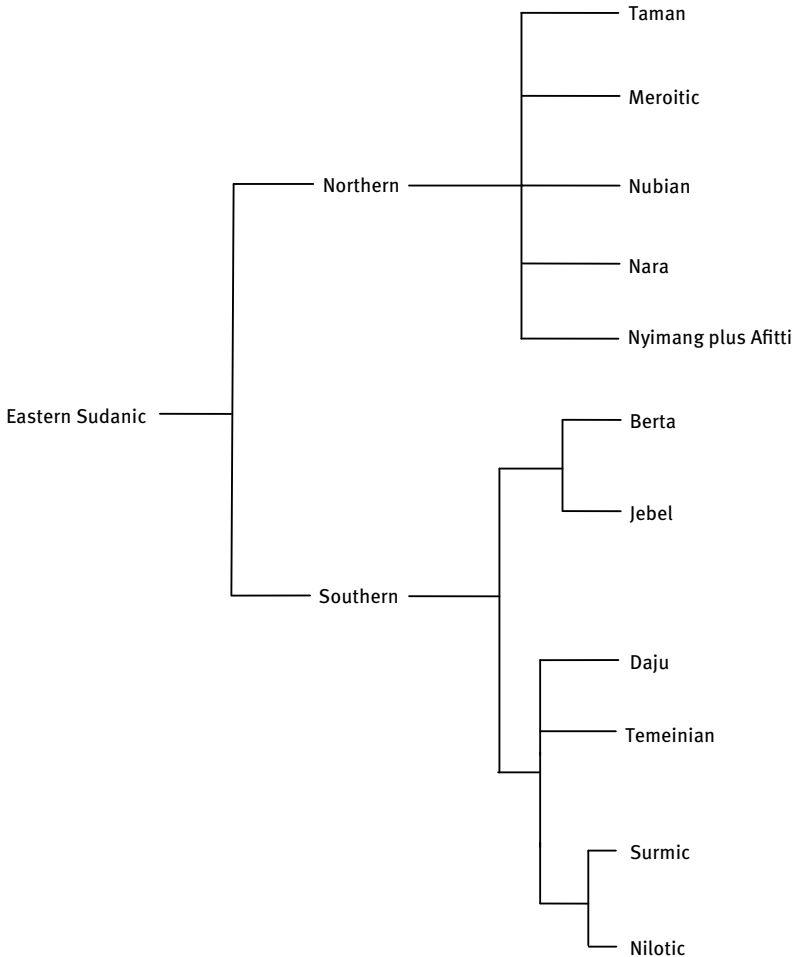
Fig. 1: The Eastern Sudanic branch within Nilo-Saharan.

Nilo-Saharan subgroup manifests typologically interesting features, this is briefly discussed below.

Karko is a verb-final language, again a feature shared with most other Northeastern Nilo-Saharan languages, except for the Southern branch of Eastern Sudanic and Kuliak, where a verb-initial or verb-second structure is common. Central Sudanic languages, on the other hand, tend to have S AUX OV order (the Eastern branch) or SVO (the Western branch). Unlike these Central Sudanic languages, Karko and most other Northeastern Nilo-Saharan languages also express grammatical relations by means of case marking rather than fixed constituent order.

Number distinctions in Karko are attested in all major lexical categories, including personal pronouns, demonstratives, interrogatives, nouns, adjectives, participles,
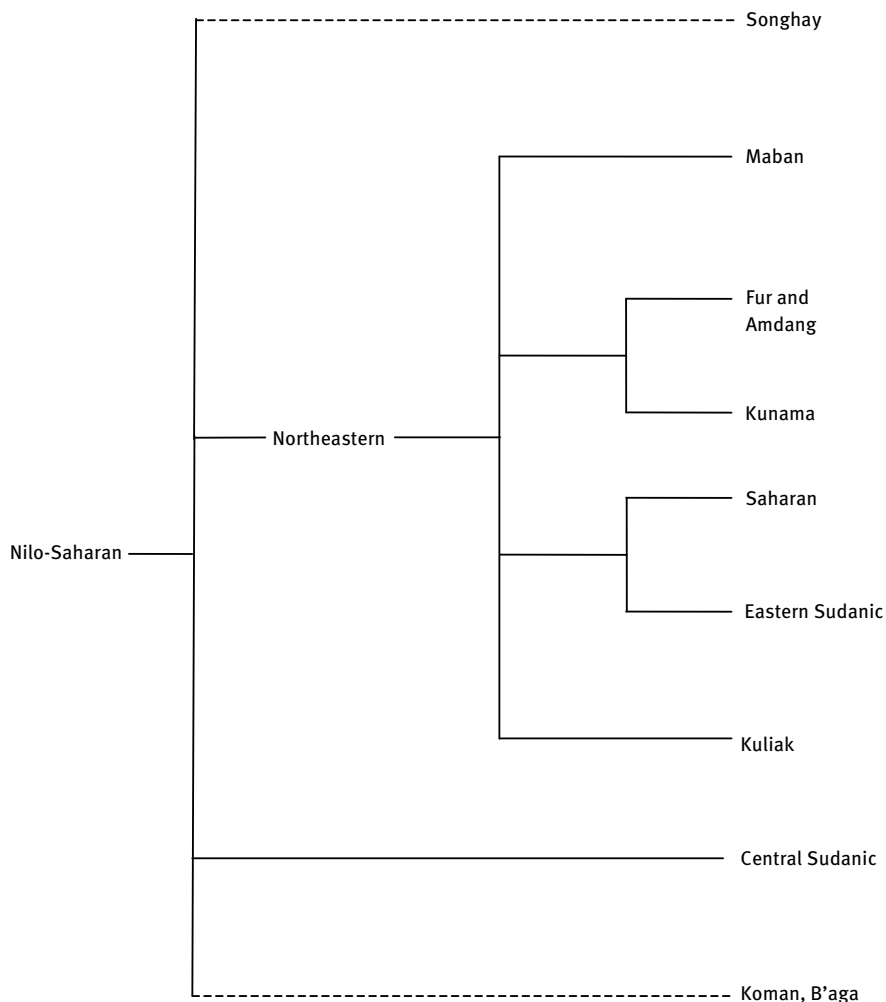
Fig. 2: The subclassification of Nilo-Saharan.

verbs, and converbs. Number values attested by nouns include (morphologically un-marked) singular, (morphologically unmarked) plural, singulative (morphologically marked, based on an unmarked plural form), plurative (morphologically marked, based on an unmarked singular form), and transnumeral (unmarked, not specified for number). Adjectives, demonstratives, participles, and pronouns show a contrast between singular and plural (or non-singular); personal pronouns additionally ex-press clusivity (including or excluding the addressee) associated with the 1st person plural. The inflectional operations are realized by suffixes, enclitics, or stem alterna-tions in Karko.

Verbal number is realized by singular or plural stems, and interacts with event number as well as with the number of S or P that are affected by the event. Agree-
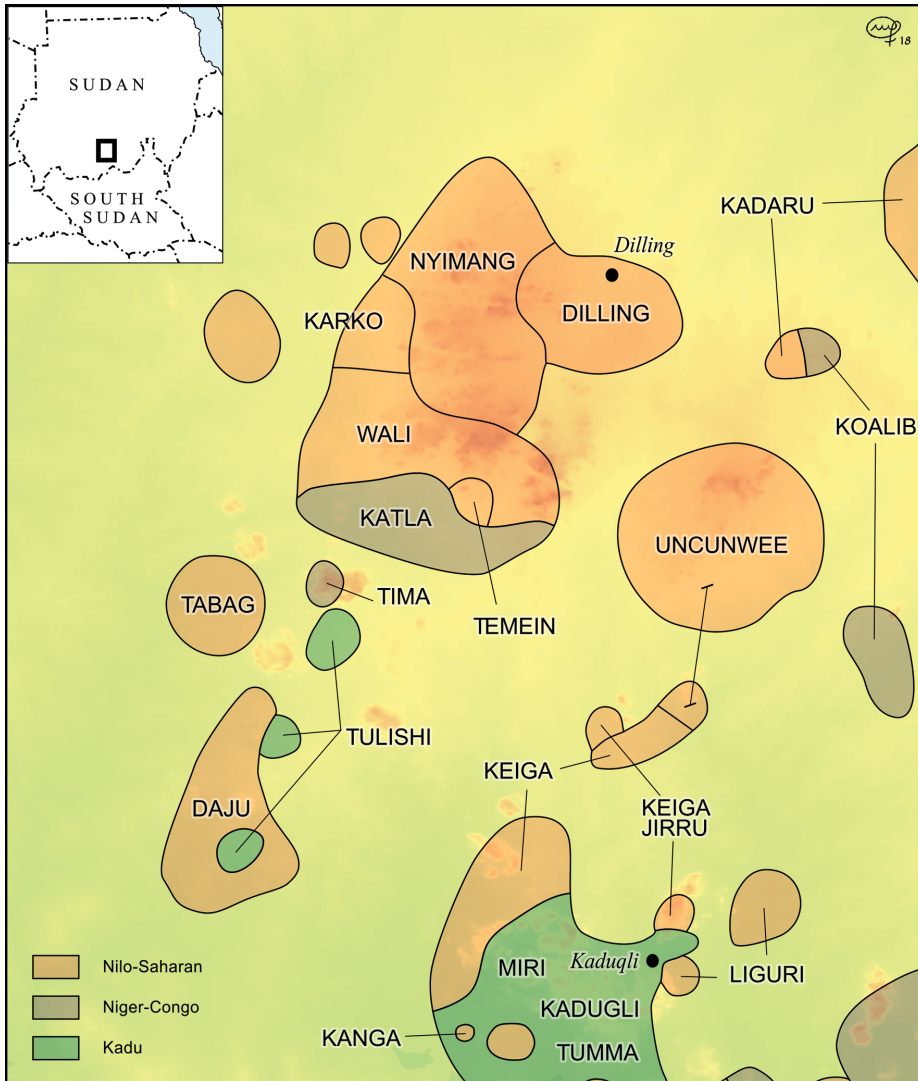
Map 2: The approximate geographical location of Karko and other Nuba Mountain languages some 70 years ago.

ment in Karko, by contrast, is realized by suffixes on the verb which cross-reference the syntactic subject. Elsewhere in Nilo-Saharan, pronominal objects are sometimes expressed on the verb as well. Only a few Nilo-Saharan languages (for example the Nilotic language Bari) do not have pronominal subject (or object) marking on verbs.

2 Pronominal, nominal, and verbal number

2.1 Generalities

Variation in number marking on major categories between Nilo-Saharan languages is in line with the typological hierarchy of morphological features proposed by Corbett (2000: 56): speaker > addressee > 3rd person > kin > human > animate > inanimate. In one of the two primary branches, Central Sudanic, number marking is expressed on pronouns, but with nouns it tends to be restricted to animate or human categories. With the former, either separate pronouns are used, or the plural forms are based on the singular, as in Moru: *má* 'I', *àmà* 'we', *mí* 'you (sg)', *àmì* 'you (pl)', *ndá* '(s)he', *ndáka* 'they' (data from Tucker and Bryan 1966: 41).

Many languages in the other primary branch of Nilo-Saharan, Northeastern Nilo-Saharan, tend to have more extensive number-marking systems for nouns, as further discussed below.

Verbal number is distinct from agreement in most Nilo-Saharan languages, being realized by singular or plural stems which interact with event number and with the number of S or P participants that are affected by the event. Agreement, by contrast, cross-references core arguments (subject and object), as in a number of Northeastern Nilo-Saharan languages discussed below. Agreement involves subject-marking affixes or proclitics, depending on the language, sometimes combined with pronominal object marking (by means of affixes or clitics). Karko employs pronominal subject suffixes but no object affixes.

2.1 Pronominal number

Pronouns inflected for number in Karko comprise the independent subject pronouns, demonstrative pronouns, and the interrogative pronouns 'who' and 'which'. As shown in Table 1, the 1st, 2nd, and 3rd person singular pronouns and also the 1st person plural inclusive pronoun (glossed as 1IN) are characterized by a short vowel. The 1st plural exclusive (glossed as 1EX) and the 2nd and 3rd person plural pronouns, by contrast, have long vowels and a low-high tone pattern, suggesting that they are formally marked for number whereas the 1st, 2nd, and 3rd person singular and the 1st person plural inclusive forms are unmarked. The fact that the 1st person plural inclusive pronoun *é* displays a short vowel like the 1st, 2nd, and 3rd person singular pronouns suggests that its referent is semantically conceived of as a single unit. This pronoun is not attested with honorific uses. When addressing a respected person or respected people, the 2nd singular or 2nd plural pronouns, *á* and *ùú*, are used.

Tab. 1: Independent subject pronouns in Karko.

1SG	*è*
2SG	*á*
3SG	*tē*
1IN	*é*
1EX	*àá*
2PL	*ùú*
3PL	*tĭ́*

Tab. 2: Independent subject pronouns and subject agreement markers in Karko.

			Transitive 'close'		Intransitive 'exist'	
		Indep subject pronouns	**Imperfective**	**Perfective**	**Imperfective**	**Perfective**
SG	1	*è*	*ɕə̄ʈə̀g-ər̀*	*ɕə̄ʈ-ngàl-r* [*ɕə̄ʈŋgàlɖ*]	*ɕwà-r*	*ɕwàà-nɟ-à-r* [*ɕwààɲɟàr*]
	2	*á*	*ɕə̄ʈə̀g-ər̀*	*ɕə̄ʈ-ngàl-r* [*ɕə̄ʈŋgàlɖ*]	*ɕwà-r*	*ɕwàà-nɟ-à-r* [*ɕwààɲɟàr*]
	3	*tē*	*ɕə̄ʈə̀g*	*ɕə̄ʈ-ngàà* [*ɕə̄ʈŋgàà*]	*ɕwàà*	*ɕwàà-nɟ-àà* [*ɕwààɲɟàà*]
PL	1IN	*é*	*ɕə̄ʈə̀g*	*ɕə̄ʈ-ngàà* [*ɕə̄ʈŋgàà*]	*ɕὲὲ*	*ɕὲὲ-nɟ-ὲὲ* [*ɕὲὲɲɟὲὲ*]
	1EX	*àá*	*ɕə̄ʈə̀g-ər̀*	*ɕə̄ʈ-ngàl-r* [*ɕə̄ʈŋgàlɖ*]	*ɕὲ-r*	*ɕὲὲ-nɟ-ὲ-r* [*ɕὲὲɲɟὲr*]
	2	*ùú*	*ɕə̄ʈə̀g-ər̀*	*ɕə̄ʈ-ngàl-r* [*ɕə̄ʈŋgàlɖ*]	*ɕὲ-r*	*ɕὲὲ-nɟ-ὲ-r* [*ɕὲὲɲɟὲr*]
	3	*tĭ́*	*ɕə̄ʈə̀g*	*ɕə̄ʈ-ngàà* [*ɕə̄ʈŋgàà*]	*ɕὲὲ*	*ɕὲὲ-nɟ-ὲὲ* [*ɕὲὲɲɟὲὲ*]

Although the independent subject pronouns distinguish between seven forms, including a distinction between 1st person plural inclusive and exclusive, the subject agreement markers exhibit fewer distinctions due to person and number syncretism. In fact, there is only a single overt subject person marker, the suffix *-(V)r* (*r* is realized as *ɖ* after *l*), which marks the 1st and 2nd person singular and the 1st exclusive and 2nd person plural, whereas the 3rd person singular and plural and the 1st person inclusive plural are unmarked, as illustrated in Table 2.

Whereas it is common for Nilo-Saharan languages (both within the Central Sudanic and the Northeastern Nilo-Saharan branch) to mark pronominal subjects on the verb by means of prefixes (or proclitics), Karko patterns along with other Nubian

languages or Eastern Sudanic languages like Berta and a number of Surmic and Nilotic languages, as well as Kuliak and Saharan, in that it only has subject suffixes (Dimmendaal et al. 2019). As shown in Table 2, the singular or plural number of the intransitive S argument is reflected both by the selection of singular or plural verb stems and by the contrasting suffixes -*à* (1st and 2nd singular) vs. -*ɛ̀* (1st plural exclusive and 2nd plural) and -*àà* (3rd singular) vs. -*ɛ̀ɛ̀* (1st plural inclusive and 3rd plural). However, the number of the transitive A argument is neither reflected by the verb stem nor indexed on the verb. The independent subject pronouns are added in Table 2 in order to show the formal contrast with the pronominal suffixes.

A number distinction in Karko interrogative pronouns is only attested for 'who' and 'which' (a pattern also found in other Nilo-Saharan languages; see the relevant sections in Tucker and Bryan 1966). Examples (1) and (2) illustrate 'who' in copula clauses and (3) and (4) in verbal clauses, the plural form of 'who' being marked by -*r*. This plural suffix is also attested in a replacive pattern, as shown in Table 4.

(1) *gɔ̄   ííd   ɖìɛ̀ɛ̀-nɖ*
     this man who-COP
     'Who is this man?'

(2) *gɔ̄ŕ   ín   ɖìɛ̀ɛ̀-r-ēnɖ*
     these people who-PL-COP
     'Who are these people?'

(3) *ɖìɛ̀ɛ̀=g   á   kɔ̄k-ngâl-r*
     who=ACC 2SG hide-TR.PFV-2.Q
     'Whom (sg) did you hide?'

(4) *ɖìɛ̀ɛ̀-r-ɛ́g   á   kɔ̄k-ɛ̄ɛ̄-ngâl-r*
     who-PL-ACC 2SG hide-PLR-TR.PFV-2.Q
     'Whom (pl) did you hide?'

The singular and plural forms of 'which' exhibit vowel alternation, which is also attested with other word categories, e.g. nouns, adjectives, and verbs, as we will show below.

(5) *ēkɕīī   nwā-ŋ-áà*
     which.SG escape-ITR.PFV-SG.Q
     'Which one has escaped?'

(6) *ēkɕɛ̄ɛ̄   nwā-k-ŋ-ɛ́ɛ̀*
     which.PL escape-PLR-ITR.PFV-PL.Q
     'Which ones have escaped?'

Tab. 3: Karko demonstratives pronouns.

Singular	Plural	Gloss
gɔ̄	gɔ́ɔ́ ~ gāŕ	'this / these'
wáŋ	wăŋ	'that / those'
tē	tĭĭ	'that / those (far away)'

The alternation of the singular and plural verb stems *kɔ̄k* vs. *kɔ̄k-ɛ̄ɛ̄* in (3) and (4) and *nwā* vs. *nwā-k* in (5) and (6) is triggered by the number of the transitive P and the intransitive S. The interaction of verbal number with participant number will be discussed in detail in Section 2.4.

Demonstrative pronouns (which also serve as determiners in Karko NPs) also manifest a singular-plural contrast. Similarly to the independent person pronouns, the plural forms are characterized by stem alternation involving a mid-high or low-high tone pattern and, at least in the case of *gɔ́ɔ́ ~ gāŕ* and *tĭĭ*, by longer segmental forms (Table 3). This suggests that these forms are marked for plural, whereas the corresponding singular forms are formally unmarked.

A functionally inclusory construction in which reference to the same entity is made twice, once by a plural pronoun and the second time explicitly, for instance by a proper name, is not attested in Karko. An inclusory meaning, however, can be expressed by coordinated NPs marked by the clitic *=ɛ̄ɛ̄r* on the first constituent and the comitative *=ɔ́ɔŋ* on the last. A personal possessive prefix on the body part term *ɖòòl* 'neck' expresses an emphatic pronoun. When such a coordinated construction represents the subject of an intransitive clause, the verb requires the 1st person plural inclusive form, as seen in (7). The same form is used when cross-referencing the 1st person plural inclusive subject pronoun *é*, as illustrated in (8).

(7) *Ahmed=ɛ̄ɛ̄r ɔn-ɖòòl=ɔ́ɔŋ   kāāk   ɕɛ́ɛ́-ŋ-ɛ̀ɛ̀*
     Ahmed=and 1SG.GEN-neck=COM Karko.LOC go.PLR-ITR.PFV-PL
     'Ahmed and I myself went to the Karko area.'

(8) *é   kāāk   ɕɛ́ɛ́-ŋ-ɛ̀ɛ̀*
     1IN Karko.LOC go.PLR-ITR.PFV-PL
     'We went to the Karko area.'

Logophoric pronouns do not occur in Karko. Within Nilo-Saharan this phenomenon is restricted (as far as present knowledge goes) to Central Sudanic and Western Nilotic languages.

2.3 Nominal number

With respect to number distinctions on nouns, Karko is characterized by a rich inventory of inflectional suffixes, comprising singulative suffixes on unmarked plural stems, plurative suffixes on unmarked singular stems, and a replacement pattern realized by pairs of singular and plural suffixes. This tripartite division is characteristic of Northeastern Nilo-Saharan languages (Dimmendaal 2000). A number of languages within this Nilo-Saharan subbranch, spoken in the geographically more peripheral zones of the family, have lost this system. This applies, for example, to Nubian languages spoken along the Nile in Sudan and Egypt, where only plural marking is found (see Tucker and Bryan 1966: 319–320).

2.3.1 Exponence

Exponence is a morphological notion referring to the number of categories that may cumulate into a single formative (Bickel and Nichols 2013). Most Karko number markers are monoexponential morphemes since they express only a single value. This is true for the frequently used pluratives *-Vŋ* and *-Vɲ* (Section 2.3.5) but also for other number markers such as the singulative *-Vt* and the plurative *-V̂ɲ*, whose selection is semantically motivated: while the suffix *-Vt* is only attested on nouns referring to individual items singled out from a group or natural set (Section 2.3.3), the suffix *-V̂ɲ* is selected by kinship terms (Section 2.3.4). On the other hand, the formatives *-áà* / *-V̂r* on agent nouns and *-Vd* / *-Vn* on participles are polyexponential morphemes, as they serve both as number markers and as derivational morphemes. Another instance of polyexponence in Karko is the fusion of a nominal suffix with the accusative case marker *-V̀g* in fast speech, as illustrated by the plurative suffix *-al* in the example below. The fused and the non-fused realization are in free variation.

(9) *tē   āt-âl   ʈwɛ̂ɛ̀=k   kùb-ngàà*
     3SG pot-PL.ACC bowl.PL=INS cover.PLR-TR.PFV
     'He covered the pots with bowls.'

(10) *tē   āt-ál=àg   ʈwɛ̂ɛ̀=k   kùb-ngàà*
     3SG pot-PL=ACC bowl.PL=INS cover.PLR-TR.PFV
     'He covered the pots with bowls.'

Multiple exponence is attested with diminutive forms of nouns in Karko. When the diminutive suffixes *-(àá)nd* (sg) and *-néè* (pl) attach to a number-marked noun stem, the resulting forms exhibit two number-marking suffixes.

(11) *ɲīl-t [ɟīlḍ]*   'tooth'   *ɲīl-ḍ-āánd*   'small tooth'
    tooth-SGT   tooth-SGT-DIM.SG

    āt-ál   'pots'   *āt-ál-néè*   'small pots'
    pot-PL   pot-PL-DIM.PL

Participles are characterized by low tones and are regularly marked by suffixes, *-Vd* in the singular and *-Vn* in the plural. Some participles demonstrate multiple exponence since they exhibit a second plural form consisting of the singular suffix *-Vd*, to which the plural suffix *-Vŋ* is added (the latter is also attested on nouns). So far, no semantic or pragmatic distinction between the plural doublets has been identified. Participles agree in number with the nouns they modify (see Section 3.1).

(12) Singular   Plural   Gloss
    èb-èd   *èb-èn, èb-èd-èŋ*   'washed'
    kàɲ-àd   *kàɲ-àn, kàɲ-àd-àŋ*   'shaved, licked, bold'
    ɕòɲ-òd   *ɕòɲ-òn, ɕòɲ-òd-òŋ*   'dried'

2.3.2 Absence of marking

The selection of certain number markers or number marking strategies is often restricted to semantically defined groups of nouns. Absence of marking, for instance, is attested with abstract nouns (including nouns referring to speech events), and nouns referring to units of time and locations, as well as natural pairs and collectives. There is no dedicated morpheme marking collectives. We use *collective* only as a semantic term when referring to entities occurring in natural groups. We avoid it as a grammatical term because, according to Corbett (2000: 117), it has been used in the linguistic literature to refer to very different phenomena. Abstract nouns are either inherently singular or inherently plural, as shown by the singular or plural forms of their modifiers, e.g. *kēē* / *kén* 'good (sg / pl)' and *úṭ* / *ǔṭ* 'big (sg / pl)'; see (13) and (14).

(13) Singular
    ḍēɟēk kēē   'good work'
    wàr úṭ   'big heat'
    òɟ úṭ   'big war'
    kám úṭ   'big hunger, famine'

(14) Plural
    wâɟ ǔṭ   'big victory'
    nééɲêè kén   'good story'
    wár kén   'good song'
    ɕìèè kén   'good language'

Place nouns (often derived from verbs) do not have morphologically distinct singular and plural forms. When they are modified by numerals with a value higher than one, these numerals do not trigger the otherwise expected plural marking on the noun modified. For doublets, as attested by *ɖùg-ùʈ ~ ɖùg-ùr,* see Section 2.3.7.

(15) *ɕɔ̀ʈ-ɔ̀ɖ*   'prison, jail'
      close-LOC.NOM

      ɖùg-ùʈ ~ ɖùg-ùr   'threshing place'
      beat-LOC.NOM

      ɖìg-ìʈ   'gathering place, gathering, meeting, crowd'
      gather-LOC.NOM

(16) *ɖùg-ùʈ*   *ʈɔ́ɟ=ɔ̄tré*   *àá*   *ɖūg-ŋ-ɛ̀-r*
      beat-LOC.NOM three=LOC 1EX beat.PLR-ITR.PFV-PL-1
      'We threshed at three threshing places.'

(17) *ɖìg-ìʈ*   *ɔ̂r*   *àá*   *kwàl-àr*
      gather-LOC.NOM two 1EX have.PLR-1EX
      'We have two gathering places.'

Some nouns referring to units of time, e.g. *ēēɟ* 'week', *ɔ̀mɔ̀t* 'month', *ɕɔ̄ɔ̄* 'year', are also unmarked for number when modified by a numeral whose value is higher than one. Other nouns, however, e.g. *ŭg* (sg) and *ùg-ùl* (pl) 'day', do take a plural marker in this context.[2] This shows that the presence or absence of number marking morphology on these nouns, while largely semantically motivated, is often also lexicalized.

(18) *tē*   *ēēɟ*   *ʈīɕ=ìg*   *hàkīīm*   *ɟìì-ŋ-àà*
      3SG week five=ACC hospital.LOC lie.SNG-ITR.PFV-SG
      'He has been in the hospital for five weeks.'

(19) *tē*   *ùg-ùl*   *tɔ́ɟ=ɔ̀g*   *wár*   *kwàl*
      3SG day-PL three=ACC there have.PLR
      'He is staying there for three days,' lit. 'he has three days there.'

Other nouns unmarked for number can be grouped into two classes, *transnumeral* nouns, as illustrated in (20), and nouns referring to natural pairs and collectives, as shown in (22). Transnumerals have only one form, i.e. they lack morphologically distinct singular and plural forms. The place nouns illustrated in (15) to (17) and the

---

2 Note that adverbials expressing periods of time require acc-marking, as attested in (18) and (19).

nouns referring to units of time addressed in the preceding paragraph are part of the class of transnumerals, too.

(20) *īt*   'louse / lice'
      tēē   'cow / cows'
      wār   'song / songs'
      ə̀ɕ   'hand / hands'
      kút   'knee / knees'

These transnumeral nouns receive either a singular or a plural interpretation depending on whether they take singular or plural modifiers, e.g. a demonstrative such as *gɔ̄* / *gɔ̄ŕ* 'this / these', which precedes the head noun, or an adjective like *kêl* / *kĕl* 'red (sg) / red (pl)', which follows the noun.

(21) *gɔ̄ īt*   'this louse'   *gɔ̄ŕ īt*   'these lice'
      tēē kêl   'red cow'   *tēē kĕl* 'red cows'

2.3.3 Singulative

Another class of nouns referring to natural pairs and collectives takes the singulative *-(V)t* suffix when referring to an individuated item singled out from a natural pair or collective, whereby the morphologically unmarked plural forms probably reflect their higher frequency of usage (as argued for Northeastern Nilo-Saharan languages in general in Dimmendaal 2000). The two stems of *ɟīl-t* / *ɟīīl* and *əl-t* / *ɔ̄ɔ̄l* are morphophonologically conditioned allomorphs.

(22) *ʈōŋ-ôt*   'mosquito'   *ʈŏŋ*   'mosquitos'
      ɟīl-t [ɟīlḍ] 'tooth'   *ɟīīl*   'teeth'
      əl-t [əlḍ]   'breast'   *ɔ̄ɔ̄l*   'breasts'
      ò̀k-ót   'bean'   *ò̀k*   'beans'
      ɛ̀ɛ́-t   'cheek'   *ɛ̀ɛ̀*   'cheeks'

When a subset of a natural group of items is modified by a numeral higher than one, the (unmarked) plural form of the noun referring to those items is required. This means that this plural form can refer both to a natural group and to items forming a subset of that group.

(23) *hākî̀ìm ò̀g   ɟīīl   ɔ̂r=ɔ̄g   ḍūɕ-ɛ̄ɛ̄-ngàl-ḍìì*
      doctor 1SG.ACC teeth two=ACC extract-PLR-TR.PFV-APPL
      'The doctor extracted two of my teeth,' lit. 'the doctor extracted two teeth for me.'

Tab. 4: The replacive number marking pattern involving *-(V)t*.

-(V)t / -(V)n	è-t	'baobab tree'	è-n	'baobab trees'
	ə̀m-ət	'moon, month'	ə̀m-ə̀n	'months'
-(V)t / -(V)l	ɕè-t	'rib'	ɕè-l	'ribs'
	kə̀m-ət	'guinea-fowl'	kə̀m-ə̀l	'guinea-fowls'
-t / -r	kwâ-t	'shoe'	kwâ-r	'shoes'
	gì-t	'tree sp.' (sg)	gì-r	'tree sp.' (pl)

The *-(V)t* suffix, which is used as a singulative marker in Karko on morphologically unmarked nouns, as in (22) above, is also attested in three replacive number-marking patterns, where it is in paradigmatic contrast with the plural suffixes *-(V)n*, *-(V)l*, and *-(V)r*. In this pattern, the marked plural forms refer to natural collectives, their *-(V)t* marked singular forms designating an item singled out from those collectives (Table 4).

2.3.4 Semantically motivated number marking

Number marking (including the absence of marking) on nouns is semantically motivated, as attested by the unmarked abstract nouns, transnumerals, and nouns referring to places and units of time (Section 2.3.2). The selection of the singulative marker is semantically determined, too, since it occurs on morphologically unmarked and marked nouns referring to natural pairs and collectives (Section 2.3.3) Another instance of the semantically determined selection of number markers is the plurative suffix *-V̂ɲ*, which is restricted to kinship terms. It is associated with a falling tone.

(24) *fâg* 'father' *fāg-âɲ* 'fathers'
    ām 'grandmother' *ām-âɲ* 'grandmothers'
    ɔnd 'husband's sibling' *ɔnd-ɔ̂ɲ* 'husband's siblings'

However, not all kinship terms take the plurative suffix *-V̂ɲ*. Some terms have singular and plural forms solely distinguished by tonal alternation.

(25) *ʈîɟ* 'maternal uncle' *ʈǐɟ* 'maternal uncles'
    êd 'in-law' *ěd* 'in-laws'
    ágàt 'grandfather' *ágǎt* 'grandfathers'

Nouns referring to paired body parts are either found among the transnumeral nouns, as attested by *ə̀ɕ* 'hand / hands' and *kút* 'knee / knees' (Section 2.3.2), or

among the nouns referring to natural collectives, e.g. *ɟíl-t / ɟīīl* 'tooth / teeth' and *ɛ̀ɛ́-t / ɛ̀ɛ̀* 'cheek / cheeks' (Section 2.3.3). Unpaired body part terms, by contrast, take the plurative suffix *-(V)nd*, which is, however, also attested on a few nouns that do not refer to body parts, such as 'fire' below.

(26) *ɕēē*   *ɕēē-nd*   'udder'
      ɕēēl   *ɕēl-ēnd*   'penis'
      māā   *māā-nd*   'hump of camel'
      ūk   *ūk-ūnd*   'fire'

Another semantically defined group which selects a specific number marker comprises mass nouns referring to liquids, substances, and granular material. Although in some languages mass nouns cannot receive plural marking, Karko mass nouns can take the plurative suffix *-(V)l*, which is also attested on other nouns such as *ǔg* (sg) / *ùg-ùl* (pl) 'day', and in a replacive pattern with the singulative (Table 4). Morphologically, these plural marked mass nouns behave like count nouns. The plural forms have a distributive sense. They can be translated as 'various sorts of X'.

(27) *ə̄mə̄nd/ə̄md-âl*   'foam'
      kwàɟ / kwāɟ-âl   'meat'
      ɕwì̀d / ɕwì̀d-îl   'sand'
      ʈēēɟ / ʈēɟ-êl   'oil'

The unmarked forms of mass nouns may be inherently singular or plural in Karko and Nilo-Saharan languages, as attested by the patterns of number agreement with modifiers. The Karko mass noun *óg* 'blood' is inherently singular, whereas *ə̂t* 'water' is inherently plural. Accordingly, these nouns are modified either by a singular or a plural modifier, as attested by the adjective *ōō / wāā* 'hot', whose singular and plural forms differ by means of stem alternation.

(28) *óg ōō*   'hot blood'
      ə̂t wāā   'hot water'

Agent nouns, too, form a semantically coherent group. They exhibit a regular replacive pattern: the singular is marked by the suffix *-áà*, the corresponding plural by the suffix *-V̂r*, with both markers having a falling tone. Agent nouns commonly designate animate beings who are characterized by habitually engaging in a specific activity. This is probably the reason why these nouns are derived from plural verb stems which are associated with habitual or generic events (see Section 2.4.4). However, in the case of 'murderer', the meaning is not restricted to a habitual activity;

rather, *tɔ̄f-áà* / *tɔ̄f-ɔ̂r* can refer both to a serial killer and to someone who has killed a single person.

(29) *gāɟ-áà*   *gāɟ-âr*   'walker, someone who walks around aimlessly'
    fɛ̄t-áà   *fɛ̄t-êr*   'hunter'
    tɔ̄f-áà   *tɔ̄f-ɔ̂r*   'murderer'

2.3.5 Productivity

Only two plurative suffixes, *-(V)ɲ* and *-(V)ŋ*, are considered to be productive, because they are used on borrowed nouns. The distribution of these suffixes is difficult to predict. In a few cases, the selection of *-(V)ŋ* is determined by a phonotactic restriction which does not allow *-(V)ɲ* to follow another palatal nasal. Rather, *-(V)ɲ* must be replaced by *-(V)ŋ*, e.g. *ɕɛ̄ɛ̄ɲɛ́ɛ̀-ɲ → [ɕɛ̄ɛ̄ɲɛ́ɛ̀ŋ]* 'tray'. The same rule is applied to one of the kinship terms, most of which are marked by *-V̂ɲ* in their plural forms (as shown in Section 2.3). Because of the restriction, *āɲ-âɲ* is realized with a *-(V)ŋ* suffix as *[āɲâŋ]* 'father's sister'.

(30) *-(V)ɲ*   *ārbíɛ̀ɛ̀* / *ārbíɛ́ɛ́-ɲ*   'car'
    kārááɕ / *kārááɕ-áɲ*   'copybook'

(31) *-(V)ŋ*   *kɪ̆t* / *kīt-ɪ̂ŋ*   'book'
    ɕɛ̄ɛ̄ɲɛ́ɛ̀ / *ɕɛ̄ɛ̄ɲɛ́ɛ̀-ŋ*   'tray'

The selection of these particular plurative suffixes, *-(V)ɲ* and *-(V)ŋ*, is probably determined by their frequent use on original Karko nouns, as illustrated by the following examples.

(32) *-(V)ɲ*   *ɕɵ̄kɵ̆l* / *ɕɵ̂kl-ɵ̂ɲ*   'gazelle'
    ɵ̀nɖ / *ɵ̄nɖ-ɵ̄ɲ*   'donkey'
    ìr / *īr-ɪ̄ɲ*   'river'

(33) *-(V)ŋ*   *ār* / *ār-âŋ*   'container made from cow dung'
    fùt / *fùt-ùŋ*   'bundle, parcel, stack, hole'
    kɵ̀r / *kɵ̀r-ɵ̀ŋ*   'shield'

The suffix *-(V)ŋ ~ -ŋ(V)* is common in other Nilo-Saharan languages, for example with independent plural pronouns in the Eastern Sudanic language Mararit (which belongs to the Taman cluster), *wa-ŋ* (1pl), *i-ŋ* (2pl), *eni-ŋ* (3pl), as well as in Maba (Weiss 2009), *má-ŋ* (1pl), *ká-ŋ* (2pl), *wá-ŋ* (3pl). In Fur (Jakobi 1990) it is frequently used as plural marker on nouns.

2.3.6 Number distinctions in adjectives

Adjectives in Karko can be divided into two groups, those inflected for number and those not inflected for number (the latter may be conceived of as transnumeral adjectives).[3] In the first group, number is mainly distinguished by tonal alternation, which may be combined with vowel alternation. With nouns, by contrast, this pattern is rare, e.g. *kwàâ* / *kwɛ̀ɛ̂* 'spear' and *ʈwàá* / *ʈwɛ́ɛ̂* 'bowl'. A few adjectives take the plurative suffix *-Vɲ*, as shown in (34). This suffix is also used on nouns, as shown in the Section 2.3.5. Singular and singulative suffixes, however, do not occur on adjectives.

(34)	Singular	Plural	Gloss
	ûr	*úr*	'black'
	ʈĩlḑ	*ʈĩlḑ*	'heavy'
	úʈ	*ŭʈ*	'big, important'
	ŋə̂ɲ	*ŋŭ̆ɲ*	'fat'
	ōō	*wāā*	'hot'
	ɕɛ̀r	*ɕēr-ɛ̄ɲ*	'short'

Most of the adjectives that are not inflected for number express physical properties. The singular vs. plural interpretation of the unmarked adjectives depends on the number of their heads (see Section 3.1).

(35)	*tōrɔ́ɕ*	'smooth'
	bādál	'weak (human)'
	bĭk	'blunt'

2.3.7 Summary of number marking strategies associated with nominals

Table 5 provides a summary of the various formal number marking strategies associated with Karko nouns, adjectives, and participles. Apart from suffixes employed in singulative, plurative, and replacive marking, stem alternation and tonal contrast can serve as sole number markers, as illustrated by some kinship terms (Section 2.3.4) and adjectives (Section 2.3.6). Table 5 also shows that there is no default number marker in Karko. As transnumerals do not exhibit overt number markers, they are not included in this table.

---

3 Whether or not adjectives exist as a separate category depends on the language. In the Nilotic language Lopit, for example, they do not constitute a separate category from stative verbs (Moodie 2019).

Tab. 5: Number markers on Karko nouns, adjectives, and participles.

		Singular	Plural	Gloss
singulative marking, plural unmarked	*-(V)t / Ø, t → [ɖ] / l__*	òk-ót	òk	'bean'
		əl-t [əlɖ]	ɔ̀ɔl	'breast'
plurative marking, singular unmarked	*-(V)nd* often, but not exclusively, on body part terms	ōō	ōr-ōnd	'head'
	-d realized as *[ɖ]*	kāmə̀l	kāmə̄l-d	'axe'
	-(V)l common on nouns referring to tools and mass nouns	āt	āt-ál	'water-pot'
		ɖūg-úʈ	ɖúg-ʈ-úl	'flail, beater'
		kwà̵	kwā̵-âl	'meat'
	-(V)r	bùt	bùt-ùr	'cat'
	-Vɲ on nouns and adjectives	àr	ār-ə̄ɲ	'rope'
		ʈòr	ʈór-óɲ	'old'
	-Vŋ on nouns and participles	kə̀r	kə̀r-ə̀ŋ	'shield'
		kàɲ-àd	kàɲ-àd-àŋ	'shaved, smooth, licked, bold'
	-V̂ɲ on kinship terms	ām	ām-âɲ	'grandmother'
replacive marking, both singular and plural marked	*-(V)t / -(V)n*	kà-t	kà-n	'field'
	-(V)t / -(V)l	kwèè-t	kwèè-l	'egg'
	-(V)t / -(V)r	kwâ-t	kwâ-r	'shoe'
	-(V)d / -(V)n attested on nouns and participles	ūr-ūd	ūr-ûn	'Faidherbia albida'
		bɛ̀r-ɛ̀d	bɛ̀r-ɛ̀n	'circumcized'
	stem alternation including tone and vowel contrast	êt	ĕt	'sibling, brother'
		ûr	úr	'black'
		ɖə̵̄	ɖɔ̵̆	'tall, long'
		ʈwà̃á	ʈwɛ̂ɛ̀	'bowl'
		hɔ̄ɔ	hâr	'tree'
	-(àá)nd / -néè diminutive markers, may attach to number marked sg and pl forms	bàg-àánd	bàg-àr-néè	'small/young he-goat'
	-áà / -V̂r replacive marking on agent nouns	tɔ̄f-áà	tɔ̄f-ɔ̂r	'murderer'

Furthermore, number distinctions can also be associated with suppletion and doublets. Suppletion is rarely attested in Karko nominal number formation. An instance of weak suppletion is represented by the irregular number marking on *wá-t* / *wā-gāl* 'enemy', involving the wellknown *-t* (Sections 2.3.3 and 2.3.4) and an unknown suffix *-gal*. The roots of the adjectives *māánd* 'small (sg)' and *ḍwáánéè* 'small (pl)', by contrast, exhibit strong suppletion. They are number-marked by the lexicalized diminutive markers *-(āá)nd* (sg) and *-néè* (pl). As for doublets, they are occasionally attested, even with the otherwise regularly number marked participles (Section 2.3.1). The doublets are not semantically differentiated.

(36) *kēd* / *kēd-êŋ ~ kě-n*   'sister's son'
      ḍùg-ùṭ ~ ḍùg-ùr   'threshing place'

2.3.8 Constructions expressing associative plural notions

Karko does not have a dedicated associative plural morpheme. Rather, the clitic conjunction *=ēēr* (allomorph *=ēē*) is attached to a personal name or kinship term. This complex element is employed in a possessive noun phrase in which the focal referent functions as the dependent and *îì* 'people' as the head, the dependent and the head being linked by the genitive marker *-n-*.[4] This construction renders the reading 'X and his/her group.'

(37) *Ahmed-ēē=n-îì*   *áámbɔ̄rɔ̄*   *tāā-ɕ-èɛ́*
      Ahmed-CO=GEN-people tomorrow   come-INTEN-PL
      'Ahmed and [his] group will come tomorrow.'

(38) *án-ēlḍ-ēē=n-îì*   *tā-m-ŋ-èɛ̀*
      2SG.GEN-wife.SG-CO=GEN-people   come-INCP-ITR.PFV-PL
      'Your wife and [her] group have come.'

If the conjunction *=ēēr* would be omitted, the resulting genitive NP, *Ahmed-n-îì* 'Ahmed's people', would still be perfectly grammatical. The genitive NPs in (37) and (38) closely resemble constructions in which the possessor is marked for plural, as in the next example.

(39) *ṭúnḍ-úl=n*   *kùlḍ*
      blind-PL=GEN   land
      'the land of the blind [people]'

---

4 It is not known to what extent the use of the possessive in associative plural constructions is typical of Nilo-Saharan.

In (40), both the focal referent and the associated group are encoded as genitive NPs with a pronominal possessor. They are conjoined by the clitic conjunction *=ēēr* and the comitative *=ɔ́ɔ̀ŋ*, which are also attested conjoining the NPs in example (7).

(40) *tē=n-ĕt=ēēr*　　　　　 *tĭ=n-īī=ɔ́ɔ̀ŋ*　　　　 *tāā-ɕ-ɛ̀ɛ́*
　　　3SG=GEN-bother.PL=and 3PL=GEN-people=COM come-INTEN-PL
　　　'His brothers and their people will come.'

2.3.9 Number marking in other Nilo-Saharan languages

Now let us compare number marking in Karko with number marking in other Nilo-Saharan languages. In several languages within Eastern Sudanic (though not in Karko), separate genitive linkers for singular as opposed to plural possessed nouns are used. This pattern is illustrated for the Nilotic language Nandi (Table 6), which is spoken in Kenya (data adapted from Creider and Creider 1989: 40).

Thus, the form *-nyʊ:* is used when the possessor 'my' is attached to a singular noun possessee, whereas *-cʊ:k* is used if the possessed noun refers to a plural. Since nouns can be inherently singular or plural in their basic form, the cognitive advantage in terms of parsing (by the hearer) is the fact that these linkers provide additional information about number. The nasal in the linker for the singular (also found in the Nandi examples below), as well as the velar stop *k* (alternating with *c*), are reflexes of archaic (and extremely stable) deictic markers in Nilo-Saharan.

The number-marking pattern illustrated in Table 5 for nouns in Karko is characteristic of a large set of Eastern Sudanic, and, in fact, Northeastern Nilo-Saharan languages. In terms of its internal complexity, number marking in Karko falls in the middle of the range within this family.

Whereas in Nilotic (i.e. Eastern Sudanic) languages like Turkana number inflection is rich but regular and predictable to a large extent (involving mora counting and the distinction between derived and non-derived nouns, as shown in Dimmendaal 1983: 210–258), it appears to be largely unpredictable in the Nilotic language Dinka. Andersen (2014: 231) identifies 83 singular-plural alternations for the Agar

Tab. 6: Nandi possessive suffixes.

	Singular possessed		Plural possessed	
	Singular	Plural	Singular	Plural
1	*-nyʊ:*	*-nya:*	*-cʊ:k*	*-ca:k*
2	*-ngu:ng*	*-ngwɒ:ng*	*-ku:k*	*-kwɒ:k*
3	*-nyɪ:*	*-nywa:n*	*-cɪ:k*	*-cwa:k*

dialect, involving vowel length, tone, vowel height, and voice quality (breathy versus creaky voice). The following examples illustrate some of these alternations (whereby the vowels without a subscript represent creaky voice, contrasting with breathy voice, marked by means of two dots below the vowel).

(41) | Singular | Plural | Gloss |
|---|---|---|
| pɨ̀ic | pɨ́ic | 'stirring stick' |
| bɨ̂ɲ | bɨ̤̂iɲ | 'chief' |
| bá̤ɲ | bậaɲ | 'cup' |
| nɔ̀oon | nóon | 'chief' |
| t̪íil | t̪jɛ̀ɛl | 'thistle' |
| wᵃ̆al | wàl | 'plant' |
| lwák | lwĕek | 'cattle byre' |
| rɔ́ɔw | rɔ̤̀ot̪ | 'hippo' |

When elicited from speakers, many nouns apparently exhibit variation in the plural, whereas no variation in the corresponding singular form of any noun is found. This difference is presumably due to a difference in frequency of use: some nouns are probably not used frequently in the plural and may therefore be liable to variation with respect to number marking. While number inflection appears to be largely unpredictable in simple native nouns, it is far more regular in derived nouns and loanwords, including other dialects of Dinka, as argued by Ladd et al. (2009: 666–668) in their analysis of the Rek dialect of Dinka. Thus, the following two examples are part of a larger set of regular nouns which are either derived or borrowed from Arabic or English.

(42) | Singular | Plural | Gloss |
|---|---|---|
| ma-jɔ̤̀ɔk | mi-jɔ̤̀ɔk | 'black bull with white chest' |
| ma-t̪ṳktɨ̤uk | ma-t̪ṳktɨ̤uuk | 'car' |

Ladd et al. (2009) further point out that, for the approximately 400 Dinka nouns investigated, 82 singular/plural combinations involving variation in length, or tone, or length combined with tone can be identified. Moreover, "native speakers are often reluctant to speculate about the plural of unattested nouns … They also appear quite willing to acknowledge when they are uncertain about the plural of specific nouns with which they are relatively unfamiliar (e.g. for male speakers, specialized cooking implements)" (Ladd et al. 2009: 668).

A number of Saharan languages as well as Fur and Temein also have a system of number prefixes. In Fur this is combined with the more widespread system of number suffixation (Jakobi 1990, Waag 2010). Examples from Waag (2010: 60–61):

(43) | Singular | Plural | Gloss |
|---|---|---|
| *dìwìl* | *dìwìl-ta* | 'thigh' |
| *d-ʊ́ʊ́* | *k-ʊ́w-tá* | 'stew' |
| *sɔɔra-ŋ* | *sɔɔra* | 'dry leftovers of porridge' |
| *d-uru-ŋ* | *k-uru* | 'bean' |

Moodie (2019) describes a feature of number marking, unique thus far in a Nilo-Saharan language, namely the *greater singular* in the Nilotic language Lopit, which uses the marked singular to denote large number, a phenomenon which has not been reported elsewhere in Nilo-Saharan.

(44) *ó-rɤ̀má   lɔ́xìdɔ̀ŋ   lòmê-tí   dè ʃáì nɤ̀ŋì*
3-find-PFV Lohidong millet-SG:ABS in tea his
'Lohidong found a grain of millet in his tea'

(45) *eí-ŋà-bàl-ú   íjòxoì   lòmê-tí*
1PL-PFV-harvest-VEN 1PL.NOM millet-SG.ABS
'We harvested so much millet!'

The interpretation of 'millet' as a singulative, for example when combined with the prepositional phrase 'in the tea', or as a greater singular, when used as a bare noun, thus depends on the context in which it occurs. In other words, where number suffixes in Lopit have their own inherent or plural meaning, the construction in which the singulative appears allows for a non-singular reference. The inverse (a plural admitting non-plural reference) does not appear to be attested in Nilo-Saharan.

Whereas Lopit is the only Northeastern Nilo-Saharan language so far with a *greater singular* morpheme, *-tí*, *greater plural* forms are more common, like the Maba suffix *-si:*. As shown by Weiss (2009), this suffix expresses larger numbers or quantities for nouns that are already inflected for plural (as with 'shoes' below); in addition, the *greater plural* occurs with inherently plural nouns taking a singulative in Maba (as with 'water').

(46) | *bara* | 'shoe' |
|---|---|
| *bar-tu:* | 'shoes' |
| *bar-tu:-si:* | 'different (pairs of) shoes' |

(47) | *ɛnji-ga* | 'a bit of water' |
|---|---|
| *ɛnji:* | 'water' |
| *ɛnji-si:* | 'many containers/quantities of water' |

2.4 Verbal number

Weiss (2009) shows for Maba that a tripartite number-marking system for nouns (with entity counting involving singulatives, pluratives, greater plurals, and a replacement pattern) does not preclude a parallel system for verbs (expressing event counting). As shown by Weiss (2009: 274), there are different markers expressing singulative actions or processes or pluractionality with respect to verbal event structure (as opposed to number associated with states, which is expressed on nouns or noun phrases in Maba); moreover, there is also a replacement pattern (depending on the verb). The fit between linguistic number and conceptual content, more specifically quantitativity, in this Northeastern Nilo-Saharan language therefore involves the question as to whether one is dealing with permanent or fluid states.

The following examples illustrate the replacement pattern.

(48) *wàsí-g t-úndúʃá-ŋ-à*
     fire-SG 3SG-light-SNG-PAST
     'She has lit the fire.'

(49) *wàsí-g t-úndúʃá-r-ì*
     fire-SG 3SG-light-PAST-PLR-DECL
     'She has lit the fires.'

In Karko, verbal number is realized by alternating singular and plural stems. These stems carry out a multiplicity of grammatical functions. Depending on the inherent meaning (Aktionsart) of a verb, they can reflect single, repeated, or distributed events (Sections 2.4.1 and 2.4.2). The selection of the stems is also sensitive to the number of the most affected participants, thus establishing grammatical relations with the intransitive subject S, the transitive object P, the ditransitive direct object T, or even the indirect object, provided that it has an Experiencer role (Section 2.4.3). Moreover, plural stems of some verbs are employed as the base of the imperfective. Plural stems are also required to encode transitivity alternations: in derived intransitive clauses plural verb stems may convey passive, facilitative, and reflexive readings.

2.4.1 Formal encoding of verbal number

The formal encoding of verbal number is associated with various strategies, including the suffixation of dedicated derivational morphemes, internal modification through vowel or tone alternation, and partial stem reduplication, as well as combinations of these means. According to the patterns of stem alternation, several morphological verb classes can be distinguished, here illustrated by imperative forms.

The abbreviations sng and plr are used to gloss verbal singular and plural stems, respectively.

One of these groups is characterized by an unmarked singular stem and a plural stem derived with *-Vr*.

(50) SNG   PLR   Gloss
     fɛ̀t   *fɛ̄t-ɛ́r*   'throw'
     kwàɲ   *kwàɲ-àr*   'build'
     gàɟ   *gāɟ-ár*   'walk'

Another group of verbs exhibits unmarked or *-Vr*-marked singular stems, the corresponding plural stems being derived with *-ʈVg*. Thus the *-Vr* / *-ʈVg* extensions realize a replacement pattern.

(51) SNG   PLR   Gloss
     bɔ̄g-ɔ́r   *bɔ̀g-ʈɔ̀g*   'forget'
     ɖwē-ɛ́r   *ɖwɛ̀ɛ̀-ʈɛ̀g*   'destroy'
     kìr   *kìr-ʈìg*   'strangle'

Some transitive verbs have an *-ɛɛr*-marked singular stem and an unmarked or *-Vk*-marked plural stem. However, with a few verbs the *-ɛɛr*-extension (allomorph *-ɛɛ*) refers to plural objects, as attested by 'extract' in (23) and 'tie' in (59).

(52) SNG   PLR   Gloss
     bɛ́ɕ-ɛ̀ɛ́r   *bɛ̀ɟ-ɛ̀k*   'tear'
     ɕîl-ɛ̀ɛ́r   *ɕìl-ìk*   'kindle fire'
     ɕwák-ɛ̀ɛ́r *ɕwàk*   'pour away'

Another group of verbs is characterized by an unmarked or *-Vr*-marked singular stem and a plural stem derived with the suffix *-ɕ*. Due to phonological restrictions, a root-final *l* or *r* is deleted before *-ɕ*. In addition to the stem extension, there may be vowel alternation, as attested by *àl* and *iɕ* 'open, take out'. The stacking of extensions is also quite common: the plural stem *bīɕír* 'scatter', for instance, exhibits two derivational morphemes, *-ɕ* and *-Vr*.

(53) SNG   PLR   Gloss
     àl   *ì-ɕ*   'open, take out'
     kìl   *kì-ɕ*   'jump over'
     əŋ-ə́r   *ɔ̀-ɕ*   'seize'
     bìr   *bī-ɕ-ír*   'scatter'

Finally, there is a group of verbs without distinct singular and plural stems, both stems being marked by the plural extension *-Vr*. The selection of this pattern ap-

pears to be semantically motivated, since all verbs exhibiting this pattern refer to social activities (e.g. greet, chat, refuse, ask, deceive, sing) affecting more than one animate (most commonly human) participant. The appearance of the verb 'sing' in this class is due to the fact that in the Karko culture singing is commonly associated with antiphony and hence with the cooperation of several participants.

(54) SNG=PLR   Gloss
    hə̄ŋ-ə́r   'greet'
    kūɕ-úr   'chat'
    wār-ár   'sing'

2.4.2 Verbs with two plural stems

Whereas most Karko verbs exhibit a singular and a plural stem, there are also verbs with one singular and two plural stems. One of the two plural stems is selected when the affected plural participant is conceived of as one group acting simultaneously, while the other plural stem is required when the affected plural participant is conceived of as individuals each acting separately. Some intransitive verbs with a triple set of stems are illustrated by the unmarked imperative forms in Table 7. The *-kVn*-marked plural stems depict a single event, the other plural stems a multiple or distributed event.

While the plural stem extension *-kVn* regularly encodes group events, multiple or distributed events can be expressed by a variety of plural stems formed by vowel alternation *(tèf* vs *tèf)* or partial stem reduplication *(àk àkàg)* or even extension by a derivational suffix such as *-ɕ*, for instance. The *-kVn* extension is restricted to imperative plural forms of intransitive verbs. Because of this restriction, it does not show up in the following indicative clauses. The stem alternations of an intransitive verb express the quantification of the S participant and the event, as illustrated by 'stand, stop'. The singular verb stem *téè* refers to a singular S participant, the plural stem *tèè* to a plural S participant and the partially reduplicated plural stem *téftéfég* to a repeated event. The reduplicated stem is non-specific as to the number of S.

Tab. 7: Intransitive verbs exhibiting a singular stem and two plural stems.

Imperative			Gloss
Singular S	**Plural S conceived of as a group**	**Plural S conceived of as individuals acting separately**	
tèf	*tèf-kèn*	*tèf*	'stand, stop'
àk	*àk-kàn*	*àk àkàg*	'sit'
bùr	*bùr-kùn*	*bù-ɕ < bùr-ɕ*	'jump down'

Tab. 8: Transitive verbs exhibiting a singular and two plural stems.

Imperative			Gloss
Singular P	Plural P conceived of as a single group, single event	Plural P conceived of as various entities, distributed event	
ʈégér	ʈɛgɛ́ɛr	ʈègèɕ	'tie'
òg	ōgór	òɕór	'call'
bàg	bɔ̀k	bɔ̀kɔ́r	'steal'

(55) ə̀nɖ   téè
    donkey stop.SNG
    'The donkey stops.'

(56) ə̄nɖ-ə̄ŋ   tèè̀
    donkey-PL stop.PLR
    'The donkeys stop.'

(57) ə̀nɖ   tɛ́ftɛ́fɛ́g
    donkey stop.DST
    'The donkey stops (several times).'

(58) ə̄nɖ-ə̄ŋ   tɛ́ftɛ́fɛ́g
    donkey-PL stop.DST
    'The donkeys stop (several times).'

Two distinct plural stems are also attested by some transitive verbs. While one plural stem expresses a single event affecting the plural P when conceived of as a single group, the other plural stem expresses a distributed or multiple event in which the plural P is conceived of as consisting of various entities, each affected individually (Table 8).

The following imperative clauses illustrate that the different verb stems are sensitive both to the number of P and to the conception of the event. The singular verb stem ʈégér interacts with a singular P participant and a single event. The plural stem ʈég-ɛ́ɛr is required when the plural P participant is conceived of as comprising a single group affected by a single action, here, a group of donkeys to be tied to a single tree. The plural stem ʈég-éɕ, by contrast, is used when a plural P is conceived of as consisting of individual entities each affected by the event, here exemplified by the donkeys to be tied to various trees.

(59) ə̀nɖ=ə̀g   hɔ̄ɔ̄=t   ʈégér
    donkey=ACC tree=LOC tie.SNG
    'Tie the donkey to the tree. '

(60) *ə̄nɖ-ə̄ɲ=ə́g   hɔ̄ɔ̄=t   ʈég-ɛ́ɛ́r*
      donkey-PL=ACC tree=LOC tie-PLR
      'Tie the donkeys to the tree.'

(61) *ə̄nɖ-ə̄ɲ=ə́g   hə̄r=ə̄t   ʈég-éɕ*
      donkey-PL=ACC tree.PL=LOC tie-DST
      'Tie the donkeys to the trees.'

The stems used in the imperative are also used in indicative clauses to express distinct events: an event affecting a singular P, an event affecting a plural P conceived of as a group, and an event in which a plural P is conceived of as comprising individuated entities each affected separately.

(62) *tŏnd   ə̀nɖ=ə̀g   hɔ̄ɔ̄=t   ʈē-ngàà*
      child donkey=ACC tree=LOC tie.SNG-TR.PFV
      'The child tied the donkey to the tree.'

(63) *tŏnd   ə̄nɖ-ə̄ɲ=ə́g   hɔ̄ɔ̄=t   ʈēg-ēē-ngàà*
      child donkey-PL=ACC tree=LOC tie-PLR-TR.PFV
      'The child tied the donkeys to the tree.'

(64) *tŏnd   ə̄nɖ-ə̄ɲ=ə́g   hə̄r=ə̄t   ʈēg-éɕ-ngàà*
      child donkey-PL=ACC tree.PL=LOC tie-DST-TR.PFV
      'The child tied the donkeys to the trees.'

A few verbs realize verbal number by means of suppletive stems (see Table 9). Except for *bə̀d* 'hit!' and *ɖùg* 'beat!', which are sensitive to event number, the choice between the stem allomorphs is conditioned either by the number of the transitive P, or the number of the intransitive S, or even – in case of the ditransitive 'give' verb – by the number of the Theme. When its Theme is singular, this donative verb exhibits two stems, *tèn* 'give me/us!' and *ʈìì* 'give him/them!' That is, the selection of the first stem is determined by a 1st person Recipient, the second stem by a 3rd person Recipient. The table displays imperative forms which generally do not exhibit dedicated morphemes distinguishing the 2nd singular from the 2nd plural form. The imperative forms *ɕwìì* and *ɕèɛ̀* 'go' are an exception to this rule because they have retained the imperative suffixes, the 2nd person singular *-i* and the 2nd person plural *-ɛ*, which are employed in some other closely related Nubian languages of the Nuba Mountains. Another deviation from the common pattern are the stems *kə̀l* 'eat! (2sg)' and *kɔ̀l* 'eat! (2pl)', which alternate for the 2nd singular and 2nd plural subject (i.e. the addressee of the order).

      Contrary to Corbett (2000: 260), who claims that "the number of verbs involved in verbal number is typically rather limited," the Karko verbs in this section indicate

Tab. 9: Karko suppletive verb stems (imperative forms).

SNG	Gloss	PLR	Gloss
kə̀l	eat (sg P, 2sg addressee)	kâm	eat (pl P)
kɔ̀l	eat (sg P, 2pl addressee)		
fūr	kill, slaughter (sg P)	tɔ̀f	kill, slaughter (pl P)
ɕwìì	go (sg S)	ɕɛ̀ɛ̀	go (pl S)
ṭèn	give (sg Theme) to 1st person sg or pl Recipient	hə̀k	give/distribute (pl Theme)
ṭìì	give (sg Theme) to 3rd person sg or pl Recipient		
bə̀d	hit	ḍùg	beat

Tab. 10: Verbal duality in Nyimang (*ṭal* 'eat'), based on Tucker and Bryan (1966: 250).

	Singular or no object	Dual object	Plural object
Singular subject	ṭal	ṭal-ɪn	ṭal-ḍi
Dual subject	ṭal-ɪn	ṭal-ḍ-ɪn	ṭal-ḍi
Plural subject	tal-ḍi	ṭal-ḍi	ṭal-ḍi

that verbal number, as realized by singular and plural stems, is attested by a large number of verbs. Only a few verbs have a single underived stem, including, for instance, *ùɕ* 'remove', *kèr* 'cut into pieces', *ɕùk* 'collect', and *ùl* 'pour out so that some liquid remains in the container.'

The large variety of formal devices used in verbal number marking suggests that the selection of specific markers is sensitive to the inherent semantics of the verb. Based on evidence from a range of unrelated languages, Dimmendaal (2014: 155) assumes that this grammatical domain is subject to a high degree of communicative dynamism, and hence to historical change or reinterpretation.

The expression of distributivity is not confined to Karko verbs; nouns, too, can express distributional notions. When mass nouns take plurative suffixes they express different types of substances, as we have shown in Section 2.3.4. Unlike verbs and nouns, Karko pronouns (independent personal pronouns, for instance) do not express distributional notions.

Verbal duality is not formally expressed in Karko, but it is attested in another Eastern Sudanic member, Nyimang (Table 10), as shown by Tucker and Bryan (1966: 250). This phenomenon cannot be analysed as agreement, because Nyimang does not have cross-reference marking for pronominal subjects on verbs, nor does it distinguish between singular and dual with independent (subject or object) pronouns.

Kutsch Lojenga (1994: 281–290) presents a description of event marking in a prototypical Central Sudanic language, Ngiti, where number marking on nouns is restricted to animate nouns. Event number as expressed by verbs involves "intransitive verbs of motion which mark singular subject versus plural subject by means of suppletive pairs ..." (Kutsch Lojenga 1994: 282), as well as (mainly) transitive verbs expressing collective and distributive plurality. Collective plurality marking occurs when a plural subject or object accompanies a singular-action verb, as in example (65). Distributive plurality is found when a plural subject or object accompanies a pluractional verb, as in (66).

(65) *ma mɨ̀ nzónzo nánzi*
     1SG SC.AUX children RSM.call:NOM1
     'I am calling the children collectively.'

(66) *ma mɨ̀ nzónzo núnzi*
     1SG SC.AUX children RSM.call:PL:NOM1
     'I am calling the children individually.'

2.4.3 Grammatical relations established by verbal number

Verbal number differs from agreement. Except for the unmarked 3rd person singular and plural and the 1st person plural inclusive, agreement is realized by subject suffixes on the verb which cross-reference the S or A argument, thus showing an accusative alignment pattern (see Section 2.2). Verbal number, by contrast, follows an ergative pattern, since it is sensitive to the number of the intransitive S and transitive P. As seen in the following examples, the number of S triggers the selection of the singular stem *ɖéè* or the plural verb stem *ɖêʈ*.

(67) *téénd ɖéè*
     girl laugh.SNG
     'The girl laughs.'

(68) *ín ɖêʈ*
     people laugh.PLR
     'The people laugh.'

In a transitive clause the selection of the singular or plural stem is sensitive to the number of P.

(69) *Alì kwàâ=g fɛ́ɛ̀*
     Ali spear.SG=ACC throw.SNG
     'Ali throws the spear.'

(70) *Alì kwèê=g   fêt*
    Ali spear.PL=ACC throw.PLR
    'Ali throws the spears.'

In ditransitive clauses, it is commonly the number of the direct object (Theme) rather than the number of the indirect object (Recipient, Beneficiary) which selects the singular or plural stem.

(71) *Alì kwàâ-g   ɔ=fɛ́ɛ́-ndìì*
    Ali spear.SG=ACC 1SG.ACC=throw.SNG-APPL
    'Ali throws the spear for me.'

(72) *Alì kwèê-g   ɔ=fɛ́t-ndìì*
    Ali spear.PL=ACC 1SG.ACC=throw.PLR-APPL
    'Ali throws the spears for me.'

However, when, in a ditransitive clause, the indirect object is assigned the role of Experiencer, it is the number of this participant which triggers the selection of the singular or plural stem. This means that in ditransitive clauses the interaction between verbal number and the direct or indirect object is sensitive to the semantic role of the participant involved.

(73) *ʈɔ̀bɔ̀lḍ   tŏnd=òg   wàr=àg àà-ngàl-ɖìì*
    baobab-seed child=ACC pain=ACC relieve.SNG-TR-PFV-APPL
    'The seed(s) of the baobab has relieved the child from pain.'

(74) *ʈɔ̀bɔ̀lḍ   tóóɲéè=g   wàr=àg íí-ngàl-ɖìì*
    baobab-seed children=ACC pain=ACC relieve.PLR-TR-PFV-APPL
    'The seed(s) of the baobab has relieved the children from pain.'

2.4.4 Plural verb stems used in the imperfective

In addition to encoding grammatical relations and participant and event plurality, the plural stems of some verbs are used as bases of the imperfective. The imperfective is morphologically unmarked and has a generic or habitual event reading. This is probably the motivation for the fact that several transitive and intransitive verbs – 'bring', 'buy', 'cause', 'close, cover', 'follow', 'enter' – exhibit a plural verb stem in the imperfective paradigm, regardless of the number of P or S. This is illustrated in Table 2 and in the examples below, where the plural stem *ɕɔ̄ʈɔ̄g* (allomorph *ɕɔ̄ʈg*) is used.

(75) è   āt=ág   ʈɔ̀ɔ̀=k   ɕə̄ʈg-ə̀r
     1SG pot=ACC gourd.SG=INS close.PLR-1
     'I (usually) close the [opening of the] pot with a gourd.'

(76) è   āt-ál=àg   ʈə̀b=ə̄k   ɕə̄ʈg-ə̀r
     1SG pot-PL=ACC gourd.PL=INS close.PLR-1
     'I (usually) close the [opening of the] pots with gourds.'

In the perfective, however, the singular P selects the singular verb stem, ɕə̄ʈ, while the plural P selects the plural stem, ɕə̄ʈə̄g.

(77) è   āt=ág   ʈɔ̀ɔ̀=k   ɕə̄ʈ-ngàl-r
     1SG pot=ACC gourd.SG=INS close.SNG-TR.PFV-1
     'I closed the pot with a gourd.'

(78) è   āt-ál=àg   ʈə̀b=ə̄k   ɕə̄ʈə̄g-ngàl-r
     1SG pot-PL=ACC gourd.PL=INS close.PLR-TR.PFV-1
     'I closed the pots with gourds.'

2.4.5 The valency-decreasing function of plural verb stems

In addition to the aspectual function of some plural stems encoding the imperfective, plural verb stems have a valency-decreasing function, as shown by their use as labile verbs in Agent- and Patient-preserving intransitive clauses, where they can have passive, reflexive, and facilitative readings (for details see Jakobi and Ibrahim 2018). This suggests that plural stems are associated with a low degree of transitivity (in the sense of Hopper and Thompson 1980).

    Let us first consider two transitive clauses. Example (79) has a singular P which selects the singular verb stem *fúr*; however, (80) has a plural P which selects the plural stem *tɔ́f* (or its allomorph *tɔ́m*) 'kill, slaughter.'

(79) ííd   ə̄gə̄d=ə̂g   fúr-āngàà
     man goat.SG=ACC slaughter.SNG-TR.PFV
     'The man slaughtered the goat.'

(80) ííd   wēē=g   tɔ́m-āngàà
     man goat.PL=ACC slaughter.PLR-TR.PFV
     'The man slaughtered the goats.'

When the number of arguments is reduced from two to one, for instance in an Agent-preserving clause whose indefinite object is deleted, the plural stem is re-

quired, as seen in (81). In the Patient-preserving clause (82) the plural stem *tɔ́f* is used to convey a passive interpretation.

(81) *ííd*   *tɔ́f-ŋ-àà*
man slaughter.PLR-ITR.PFV-SG
'The man slaughtered.'

(82) * əgəd*   (*íìd=ìk=né*)   *tɔ́f-ŋ-àà*
goat.SG (man=INS=FOC) slaughter.PLR-ITR.PFV-SG
'The goat was slaughtered (by the man).'

Examples (83) and (84) are transitive clauses. The number of the P participant triggers the selection of the singular or plural verb stem. Depending on the semantic properties of the verb and the syntactic subject, a plural stem in a non-basic intransitive clause can have either a passive or a reflexive interpretation. The latter is common with verbs of body care, such as 'wash', 'dress', or 'shave', whose syntactic subject refers to an agentive human being, as in (85). A non-agentive S, such as *tèèl* 'hair' in (86), excludes a reflexive reading. Rather, the example has a passive interpretation.

(83) *tē* *tŏnd=òg*   *kèè*
3SG child.SG=ACC shave.SNG
'He shaves the (head of the) child.'

(84) *tē* *tóóŋéè=g*   *kâŋ*
3SG children.PL=ACC shave.PLR
'He shaves the (heads of the) children.'

(85) *tē* *kâŋ*
3SG shave.PLR
'He shaves (himself).'

(86) *tèèl* *kâŋ*
hair shave.PLR
'The hair is shaved.'

Plural verb stems are commonly used in transitivity alternations of the causative/inchoative type (in the sense of Haspelmath 1993), as attested by the labile verb 'shave' in the preceding examples. The verb 'break' is also a labile verb. In the transitive clauses below, the verb is associated with an Agent who causes the event and an affected Patient. In this context the singular object interacts with the inflected singular verb stem *féè* and the plural object with the inflected plural verb stem *hɔ̂ɟ*.

(87) tē   āt=ág   féè
      3SG pot.SG=ACC break.SNG
      'He breaks the pot.'

(88) tē   āt-âl   hɔ̂ɟ
      3SG pot-PL.ACC break.PLR
      'He breaks the pots.'

The plural verb stem *hɔ̂ɟ*, which is used in the preceding transitive clause, is also required in Patient-preserving constructions in which the event is conceived of as occurring spontaneously without an external Agent. In Karko, such clauses additionally express a property characteristic of the referent of the syntactic subject.

(89) āt   hɔ̂ɟ
      pot break.PLR
      'The pot breaks', 'a pot can break, a pot is fragile'

(90) āt-ál   hɔ̂ɟ
      pot-PL break.PLR
      'The pots break', 'pots can break, pots are fragile.'

3 Agreement and the syntax of number

Karko and many other Northeastern Nilo-Saharan languages are head-final at the clausal level, meaning that the verb tends to occur in final position; languages belonging to the Southern branch of Eastern Sudanic deviate from this pattern in that they tend to put the verb in initial or second position. Within the Central Sudanic branch of Nilo-Saharan one finds either SVO constituent order (in the Western branch) or a constituent order SAUXOV alternating with SVO (in the Eastern branch), as pointed out in Section 1. While the position of the verb thus tends to vary, there is a rather permanent tendency for noun phrases to be head-initial regardless of constituent order at the clausal level. For further details, the interested reader is referred to Dimmendaal et al. (2019). The formal expression of agreement within the noun phrase differs across the phylum. Below, we therefore concentrate primarily on Karko.

3.1 Number agreement within the noun phrase

Let us first consider agreement between adjectives, participles, determiners, and numerals and their head nouns. There are two groups of adjectives: i) adjectives

that have distinct singular and plural forms (mostly realized by tonal alternation, as well as, more rarely, by a plural suffix) – see (91); and ii) adjectives that have only a single form – see (92) below.

(91) Singular   Plural   Gloss

bɔ̀g	*ûr*	*bɔ̀g-ɔ̀r*	*úr*	'black he-goat'
he-goat	black.SG	he-goat-PL	black.PL	
kūg	*ʈîlɖ*	*kūg-ūr*	*ʈǐlɖ*	'heavy rock'
rock	heavy.SG	rock-PL	heavy.PL	
ííd	*úʈ*	*ín*	*ǔʈ*	'big, important person'
person	big.SG	people	big.PL	
ēlɖ	*ɕɛ̀r*	*ēl*	*ɕɛ̄r-ɛ̄ŋ*	'short woman'
woman	short.SG	women	short-PL	

Adjectives with a single form (Section 2.3.6) can have a singular or plural interpretation, depending on the number value of the noun they modify.

(92) Singular   Plural   Gloss

kɔ̄kɔ́nɖ tɔ̄rɔ́ɕ	*kɔ́kɔ́r tɔ̄rɔ́ɕ*	'smooth stone(s)'
ɕɔ̌lɖ bādál	*ɕɔ̌lɖ bādál*	'weak young man'
kɔ́nɖ bǐk	*kɔ́nɖɔ̌l bǐk*	'blunt knife'

Participles, too, agree in number with the head noun they modify. They are regularly inflected for singular and plural, as we have shown in Section 2.3.1. Some participles exhibit doublets in their plural form, as illustrated by *ɕɔ̀ɲ-òn* and *ɕɔ̀ɲ-òd-òŋ*.

(93)   *āt ɕɔ̀ɲ-òd*   'dried pot'
       āt-ál ɕɔ̀ɲ-òn ~ āt-ál ɕɔ̀ɲ-òd-òŋ   'dried pots'

Determiners (which are formally identical to independent demonstrative pronouns) precede the head noun, while other modifiers follow it. Determiners commonly agree in number with their head noun. However, when the referent of the head noun is conceived of as a single collective, it selects a singular determiner, as in (95). Whereas, in (94), the children are distributed over several houses, in (95) the children are inside one house and therefore conceived of as a single group. This conception is reflected by the selection of the singular determiner for the plural head noun. However, this distinction is not reflected in the verb, which is sensitive to the plural P and appears with the plural stem extension *-ɛ́r* in both cases.

(94) *gɔ̄ŕ*   *tóóɲē*   *kɔ̂l=ɔ̄t*   *ɕɛ̄ɛ̂=g*   *gēl-ɛ́r*
      DEM.PL children house.PL=LOC be.inside.PL=ACC look.at-PLR
      'Look at these children inside the houses!'

(95) gɔ̄   tóóɲē   kɔ̄ɔ̄l=ɔ́t   ɕɛ̄ɛ́ɛ́=g   gēl-ɛ́r
     DEM.SG children house.SG=LOC be.inside.PL=ACC look.at-PLR
     'Look at these children inside the house!'

Karko numerals do not participate in number agreement patterns when used as modifiers of NPs. Except when modifying unmarked nouns referring to collectives and transnumerals, numerals higher than one require the plural form of the head noun. Numerals immediately follow the noun they modify, but precede other modifiers, e.g. a color adjective as in (96), or an adjective expressing age as in (97).

(96) àá tēē ʈə̂ɟ kēl=ég   kwàl-àr
     1EX cow three red.PL=ACC have.PLR-1
     'We have three red cows.'

(97) gə̄ŕ tēē ʈə̂ɟ   ʈōr-ɲ=óg   ɕàɕ
     these cow three old-PL=ACC sell/buy.PLR.IMP
     'Sell/buy these three old cows.'

When a head noun takes a diminutive marker, its modifiers (for example, an adjective and a quantifier) are required to be diminutive-marked as well. This pattern of distribution is reminiscent of languages with a noun class system where modifiers must reflect the class of their head noun. The examples also show that a transnumeral noun like *tee* 'cow' and a mass noun like *ééɟ* 'milk' can take the diminutive marker, which in this context expresses a small quantity (see also Section 4).

(98) àá tēē-né   úr-nē   fáánéè=g   kwàl-àr
     1EX cow-DIM.PL black-DIM.PL few.DIM.PL=ACC have.PLR-1
     'We have a few black cows.'

(99) ééɲē   ɕír-nē   ŋwēē-néè=g   ɔ̀=ʈì-ŋgàl-ɖìì
     milk.DIM.PL dirty-DIM.PL amount-DIM.PL=ACC 1SG.ACC=give-TR.PFV-APPL
     'He gave me a small amount of dirty milk.'

3.2 Clausal agreement

Clausal agreement in Karko comprises two strategies: i) subject agreement marking on the verb, which shows the accusative alignment pattern S=A≠P, and ii) interaction between participant number and verbal number. The selection of a singular or plural stem is sensitive to the number of the intransitive S and the transitive P participant (but not to the number of the transitive A). This pattern thus reflects a

Tab. 11: Subject agreement markers on the verb.

	Imperfective	Perfective	
	Transitive and intransitive	**Transitive**	**Intransitive**
1SG	*-(V)r*	*-ngàl-r [ŋgàlḍ]*	*-nɟ-â-r [nɟâr]*
2SG	*-(V)r*	*-ngàl-r [ŋgàlḍ]*	*-nɟ-â-r [nɟâr]*
3SG		*-ngàà [ŋgàà]*	*-nɟ-àà [nɟàà]*
1IN		*-ngàà [ŋgàà]*	*-nɟ-èè [nɟèè]*
1EX	*-(V)r*	*-ngàl-r [ŋgàlḍ]*	*-nɟ-è-r [nɟèr]*
2PL	*-(V)r*	*-ngàl-r [ŋgàlḍ]*	*-nɟ-è-r [nɟèr]*
3PL		*-ngàà [ŋgàà]*	*-nɟ-èè [nɟèè]*

system of grammatical relations that exhibits ergative-like features, S=P≠A, as shown in Section 2.4.3.

In the perfective, the sensitivity of the finite verb to the number of the intransitive S is additionally reflected by the selection of vocalic suffixes which precede the subject agreement markers, as seen in Table 11. The suffix *-a* cross-references a singular S, while the suffix *-ɛ* cross-references a plural S. These suffixes combine with the overt subject agreement marker *-r* so that the 1st or 2nd person singular subjects are marked with *-à-r*, whereas the 1st person plural exclusive and 2nd person plural subjects are marked with *-è-r*. The 3rd person singular subject is cross-referenced with *-aa*, the 1st plural inclusive and 3rd plural subject by *-èè*. These number marking suffixes cross-reference the number of the intransitive S in the perfective forms; the number of the transitive A or P, however, is not cross-referenced by such suffixes.

When, in an intransitive clause, the subject refers to a mass noun, its inherent singular or plural number value is reflected on the verb by the subject agreement markers singular *-àà* and plural *-èè* and the corresponding forms of the modifiers. The low degree of transitivity, by contrast is signaled by the plural verb stem, which operates independently of subject agreement marking.

(100) *ṭēēɟ ōō ɕwák-ăm-nɟ-àà*
　　oil　hot.SG　pour.away.PLR-INCP-ITR.PFV-SG
　　'The hot oil has been poured away.'

(101) *ə́t wāā ɕwák-ăm-nɟ-èè*
　　water　hot.PL　pour.away.PLR-INCP-ITR.PFV-PL
　　'The hot water has been poured away.'

Unlike main verbs, converbs are dependent verbs with no or reduced inflectional morphology. They are followed by a clause-final main verb carrying all the inflectional properties which have scope over the preceding converb(s). Converbs are

common in Northeastern Nilo-Saharan (except for the Southern branch of Eastern Sudanic), as illustrated by data from Karko.

3.3 Converbs

Karko has two different converbs, same subject converb (ssc) and different subject converb (dsc), used to place an event temporarily in relation to the main verb and to other events expressed by other converbs in a chain. While the verb stem of a dsc is solely marked by tone, the verb stem of a ssc is morphologically marked by tone and *-n-ɖVt,* a suffix linked to the verb stem by *-n.* Depending on the final segment of the verb stem, *-n-ɖVt* has several allomorphs. Neither the ssc nor the dsc exhibits any inflectional morphemes. The inflectional values of the main verb have scope over the preceding converb(s). Like an inflected main verb, an intransitive ssc interacts with the number of S by selecting a singular stem, as attested by *nɔ́ɔ́-n-ɟɔ̄t* in (102), or a plural stem, as attested by *nɔ̄ɔ̄-n-ɟɔ̄t* in (103). The singular and plural stems of the verb 'climb' are solely distinguished by tone. The main verbs, *ɟèɛ̀-ɲ-àà* and *ɟèk-ɲ-ɛ̀ɛ̀,* interact with the number of S, too. Since they are inflected for the 3rd person perfective, these values are adopted by the preceding ssc without, however, exhibiting any inflectional morphemes.

(102) *ɖɔ̀ŋɔ̀ʈ=ɔ̄t nɔ́ɔ́-n-ɟɔ̄t   ɟèɛ̀-ɲ-àà*
   bed=LOC climb.SNG-LK-SSC sleep.SNG-ITR.PFV-SG
   'He climbed on the bed and fell asleep.'

(103) *ɖɔ̀ŋɔ̀ʈ=ɔ̄t nɔ̄ɔ̄-n-ɟɔ̄t   ɟèk-ɲ-ɛ̀ɛ̀*
   bed=LOC climb.PLR-LK-SSC sleep.PLR-ITR.PFV-PL
   'They climbed on the bed and fell asleep.'

When the main verb is inflected for the (morphologically unmarked) imperative, the ssc adopts this value. The examples below also show that the pronominal 3rd person P is not overtly expressed. However, the reference to the singular or plural P is expressed by the singular stem of the ssc *bɔ̄ɔ̄-n-dɔ́t* and the plural verb stem of the ssc *bɔ̀d-ǹ-dɔ̄t.* The number of P also selects the singular or plural stem of the main verb 'break'.

(104) *bɔ̄ɔ̄-n-dɔ́t   ŋɔ̄ɔ̄ŋɛ̄ɛ́r*
   hit.SNG-LK-SSC break.SNG
   'Hit and break it!'

(105) *bɔ̀d-ǹ-dɔ̄t   ŋɔ̀ŋ*
   hit.PLR-LK-SSC break.PLR
   'Hit and break them!'

In a sequence of events with different subjects the non-final verb is realized as dsc. This is illustrated by the verb 'cause' in the periphrastic causative clauses below. The subject of the transitive verb 'cause' is *kēēd* 'cold'. It differs from the pronominal subject of the main clauses, *è* expressing the 1st person singular and *é* expressing the 1st person plural inclusive. Whereas the singular verb stem *ə̌t* is selected by the acc-marked 1st person singular P *ɔ̀g*, the plural stem *ɔ̌t* is selected by the acc-marked 1st person plural inclusive P *íĩg*. The stems of the intransitive main verb, *ɟìi* and *ɟèè*, are selected by the singular and plural S, respectively. The singular S is additionally cross-referenced with the inflectional suffix *-à-r*, while the plural S is cross-referenced with *-èè*.

(106) *kēēd ɔ̀g ə̌t è kə́lḍ ɟìi-ŋ-à-r*
 cold 1SG.ACC cause.SNG.DSC 1SG house.LOC sleep.SNG-ITR.PFV-SG-1
 'The cold made me sleep inside the house.'

(107) *kēēd íĩg ɔ̌t é kə́lḍ ɟèè-ŋ-èè*
 cold 1IN.ACC cause.PLR.DSC 1IN house.LOC sleep.PLR-ITR.PFV-PL
 'The cold made us sleep inside the house.'

4 Semantics and discourse

In Karko and many other Northeastern Nilo-Saharan languages, plural nouns referring to natural pairs and collectives are often morphologically unmarked. Items singled out from these pairs or groups receive singulative marking. Morphologically unmarked plural forms probably also reflect a higher frequency of usage, as argued in Dimmendaal (2000), with nouns referring to pairs, animals usually living in larger groups, or objects usually occurring in larger numbers tending to belong to this category.

The following examples first illustrate the unmarked plural form *ɟīl* 'teeth'. The marked singulative form *ɟīl-ḍ* 'tooth', by contrast, is used when referring to a single tooth, see also (22).

(108) *ɟīīl=íg kɔ̄g-ṭɔ̄r-ɔ̂r*
 tooth=ACC clean-DIR-PLR
 'He cleans [his] teeth frequently.'

(109) *ííd gɔ̄ kām=ág ɟīl-ḍ=ík ɕēē-ŋám*
 person this food=ACC tooth-SGT=INS pierce-TR.PFV.NEG
 'Nobody has touched this food', lit. 'person did not pierce this food with a [single] tooth.'

This pattern is also attested in Cushitic or Semitic languages and appears to be part of a set of areal characteristics, which include extensive case marking, a verb-final syntax, converbs, and other features (Dimmendaal 2008). While singulative marking is not attested in many languages, one can expect the frequency of some relevant nouns to be lower in the singular. Whereas most Indo-European languages (with the exception of Celtic languages) do not have singulative marking, there is evidence, for example from Italian, that some nouns tend to be used more frequently in the plural than in the singular. Baayen et al. (1997) carried out a psycholinguistic experiment involving visual word recognition on a screen with speakers of Italian (a language with singular versus plural suffixation, as in *nas-o/nas-i* 'nose' or *dent-e/dent-i* 'tooth'). Speakers were asked to decide as quickly as possible whether a letter string presented on a screen was an Italian word or not. Interestingly, some nouns were recognized faster when a plural form was shown on the screen than when a singular form was shown, as was the case with *dent-i* 'teeth' as against *dent-e* 'tooth', whereas with nouns such as *nas-o* versus *nas-i* it was the other way around. In psycholinguistics, frequency of usage is known to be a reliable predictor of response time. Nouns belonging to the so-called plural dominant category, such as 'tooth/teeth', prototypically involve singulative marking in Northeastern Nilo-Saharan languages. Hence, these Nilo-Saharan languages "... represent a natural correlation between values of different grammatical parameters – namely, functional as opposed to formal markedness" (Dimmendaal 2000: 254).

In the distantly related Nilo-Saharan language Maba (Weiss 2009), which belongs to the Maban branch within Northeastern Nilo-Saharan, one also finds the "classical" Northeastern Nilo-Saharan tripartite number marking system, with singulatives, pluratives and a replacement pattern (for example *súmbúrì-g* (sg) / *súmbúr* (pl) 'twin(s)'; *tàndàm* (sg) / *tàndàm-íː* (pl) 'giant kudu(s)'; *kámbà-g* (sg) / *kàmbáɲì:* (pl) 'boy(s)'. However, this language manifests additional variation on the theme, namely the expression of quantification either on the noun or on the verb.

Transnumeral interpretation in Maba occurs only with inanimate nouns; with other nouns, the explicit marking of number on modifiers (adjectives etc.) shows the intended interpretation:

(110) *kɔ́drɔ́   sɔ̀llɔ́kɔ́*
    stone(s) slippery
    1. 'slippery stones'; 2. 'a slippery stone'

    íríː   kùllà-g
    leopard big-sɢ
    'big leopard' (plural: *kúllέy*)

In Maba, number values expressed by nouns comprise singular, plural, singulative, plurative, greater plural, and associative plural. Some plural nouns can be assigned to two morphologically defined groups: i) unmarked nouns referring to natural pairs

or collectives, which take a singulative suffix when referring to an individuated item; and ii) transnumeral nouns having a single form, which, depending on the singular vs. plural form of the modifier, may have a singular or plural interpretation.

There is no dedicated paucal marker in Karko (or other Nilo-Saharan languages, as far as present knowledge goes). However, when the Karko diminutive *-néè* (pl) attaches to the marked or unmarked plural form of a mass noun, it triggers a paucal interpretation. As we have shown in Section 2.3.4, plural-marked mass nouns have distributive connotations. Thus, *ɕwììd-îl* may be translated as 'various kinds of sand'. When such a plural-marked mass noun takes the diminutive *–néè* (pl), the resulting *ɕwììd-îl-néè* is still associated with a distributive connotation and may be translated as 'various small quantities of sand'.

(111) Mass noun	Mass noun + Diminutive plural	Gloss
wèè	*wèè-néè ~ ɔ̀ɔ̀ɲéè*	'small quantity of sorghum'
ɔ̄k	*ɔ̄k-néè*	'small quantity of urine'
tîr	*tír-néè*	'small quantity of beer'
ɕwììd-îl	*ɕwììd-îl-néè*	'various small quantities of sand'

However, when a mass noun takes the diminutive *-(àá)nd* (sg), it expresses a small portion or part of the substance referred to. This suffix tends to fuse with the noun.

(112) Mass noun	Mass noun + Diminutive singular	Gloss
ɕwììd	*ɕwīīnd*	'grain of sand'
mānd	*mɔ̄ɔ́nd*	'piece of grass'
óg	*óóŋg*	'drop of blood'

The Karko system of classification of nouns according to their pattern of expressing number distinctions, therefore, differs from the system in noun class languages, where modifiers agree with the noun class of the head noun (Dimmendaal 2000: 214).

Karko is characteristic of Nilo-Saharan languages in general in that semantically defined groups of nouns, such as kinship terms or derived nouns expressing agents, instruments, or locations select their own specific number-marking suffixes.

Mass versus count nouns do not behave uniformly in terms of individual lexical items in Northeastern Nilo-Saharan. In some languages mass nouns behave as non-alternating nouns, and in others as a plural-singulative class – but similarly, within each language they may be inflected for number or not, as in the Nilotic language Bari, where *lɛ* 'milk' may take a singulative marker *lɛ-tat* 'a drop of milk', but 'ashes' *kurök* (or *kuröŋ*) does not. In the Maba language, which also retained the "classical" tripartite system of number inflection like Bari, the noun 'ashes' *ɔ̀wún*, may take a

singulative marker *ɔ̀wùnú-g* 'a bit of ashes', but the word for 'milk' does not. Singulatives may also express diminutive meaning, as with *kíjí:* 'head', *kìjí-g* 'small head' in Maba. The semantic range covered by singulative markers in Maba again does not appear to be arbitrary. In contrast to words like 'ashes' (which can be divided into parts without losing any of their defining properties), a word like 'head' designates a minimal concrete entity which would lose some of its defining properties if divided into parts. Consequently, the singulative here expresses a 'smaller variety of'.

Numeral classifiers as such do not occur in Karko or other Nilo-Saharan languages. However, there are collective classifiers which are commonly employed as heads of possessive noun phrases. The items conceived of as collectives are employed as dependents, the head and the preceding dependent being linked by the genitive marker *-n-*. The dependent is in a plural form, as attested by *wèè* and *gīɕīl* (unmarked for plural). The noun *tēē* is a transnumeral noun (unmarked for number).

(113) *ɕùkùd*   'heap'   *wèèŋùkùd*   'heap of sorghum'
       fût   'bundle, stack'   *gīɕīlm̄bùt*   'stack of money'
       ɗwātâd   'herd'   *tēēnɗwātâd*   'herd of cattle'

The most detailed account of number marking on nouns from a semantic perspective so far is to be found in Storch (2005: 380–416) in her analysis of the noun morphology of the Western Nilotic branch of Nilotic. The author presents a discussion of semantic notions associated with singular and plural suffixes, but also of prefixes expressing natural gender, as well as other derivational notions. The author also relates some of these grammatical features to phenomena relevant in the culture of speakers of these languages.

There are no obvious cases of discourse-sensitive number marking in Karko or other Northeastern Nilo-Saharan languages. However, these languages sometimes do manifest phenomena which are best explained as construction-type effects, such as the greater singular in Lopit, as discussed in Section 2.3.9.

5 Conclusions

In Karko, linguistic number marking is carried out by means of several morphological and lexical devices which serve to express quantification. These devices include the absence of number markers on nouns referring to natural pairs or collectives. The lack of number markers on such nouns contrasts with singulative marking, which is used to single out individual items from such natural pairs and collectives. Similarly, the use of diminutive markers on mass nouns is a morphological device to express small quantities of the substances referred to by these nouns. Finally, the

distinction of singular and plural verb stems is – except in the case of suppletion (see Table 9) – a non-segmental morphological device for quantification. This is particularly obvious when a verb has two plural stems, one referring to an event affecting a single group of participants, and another referring to a multiple event distributed in space or time. This means that there is a fairly close fit in general between linguistic number marking and conceptual content. Nevertheless, interesting deviations occur, as with the greater singular in Nilotic languages like Lopit.

Number as a category is not relevant to discourse structure in Karko, although there are construction-type effects, as illustrated in Section 2.3.9 above. Its general nature as a lexical, morphosyntactic, and semantic property of languages belonging to this phylum is reasonably well understood.

The most notable aspect of the number category with verbs in Karko is the fact that – in addition to quantifying the number of an event or the number of affected participants in an event – plural verb stems are required in non-basic intransitive clauses. In Agent-preserving and Patient-preserving constructions, plural verb stems can have a passive, reflexive, or facilitative interpretation, depending on the semantics of the verb and its arguments. This suggests that plural verb stems are associated with a low degree of transitivity (in the sense of Hopper and Thompson 1980). That is, plural verb stems have adopted new grammatical functions in the domain of valency-reducing devices.

As detailed descriptions are still lacking for language groups such as Temeinian and Maban languages other than Maba, or languages like Nara for that matter, data from these Nilo-Saharan members may affect the general picture sketched above. Furthermore, the historical links between the various systems found in different branches or individual languages are still poorly understood. This applies in particular to the origin of nominal number marking by means of prefixation in Fur or Temein languages, as opposed to the common suffixation strategy found elsewhere in the phylum, but also the question of the extent to which extensive number-marking systems are memorized or are derived by productive lexical rules.

Acknowledgements

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
A	Agent
ABS	absolutive
ACC	accusative
APPL	applicative
AUX	auxiliary
CO	conjunction
COM	comitative
COP	copula
DEM	demonstrative
DECL	declarative
DIR	directive applicative
DSC	different subject converb
DST	distributed
EX	exclusive
FOC	focus
GEN	genitive
IMP	imperative
IN	inclusive
INCP	inceptive
INS	instrumental
INTEN	intentional
ITR	intransitive
LOC	locative
NEG	negation
NOM	nominative (when occurring in noun phrases), nominalizer (when occurring on verbs)
P	Patient
PFV	perfective
PL	plural
PLR	verbal plural
RSM	resumptive marker
S	intransitive Subject
SC	subject concord
SG	singular
SGT	singulative
SNG	singular verb stem
SSC	same subject converb
TR	transitive
VEN	ventive.

References

Andersen, Torben. 2014. Number in Dinka. In Anne Storch & Gerrit J. Dimmendaal (eds.), *Number – Constructions and Semantics. Case studies from Africa, Amazonia, India and Oceania,* 221–263. Amsterdam/Philadelphia: John Benjamins.

Baayen, Harald, Cristina Burani & Robert Schreuder. 1997. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1996*, 13–33. Dordrecht: Kluwer Academic Publishers.

Bickel, Balthasar & Johanna Nichols. 2013. Exponence of selected inflectional formatives. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/21, Accessed on 2021-02-08.)

Corbett, Greville. 2000. *Number*. Oxford: Oxford University Press.

Creider, Chet A. & Jane Tapsubei Creider. 1989. *A Grammar of Nandi*. Hamburg: Buske.

Dimmendaal, Gerrit J. 1983. *The Turkana Language*. Dordrecht: Foris Publications.

Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan Languages. *Anthropological Linguistics* 42(2). 214–261.

Dimmendaal, Gerrit J. 2008. Africa's verb-final languages. In Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa,* 271–307. Cambridge: Cambridge University Press.

Dimmendaal, Gerrit J. 2014. Pluractionality and the distribution of number marking across categories. In Anne Storch & Gerrit J. Dimmendaal (eds.), *Number – Constructions and Semantics. Case studies from Africa, Amazonia, India and Oceania*, 57–75. Amsterdam/Philadelphia: John Benjamins.

Dimmendaal, Gerrit J., Colleen Ahland, Angelika Jakobi & Constance Kutsch Lojenga. 2019. Linguistic features and typologies in languages commonly referred to as 'Nilo-Saharan'. In Ekkehard Wolff (ed.), *Cambridge Handbook of African Linguistics,* 326–381. Cambridge: Cambridge University Press.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Maria Polinsky & Bernard Comrie (eds.), *Causatives and Transitivity,* 87–120. Amsterdam: John Benjamins.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Jakobi, Angelika. 1990. *A Fur Grammar*. Hamburg: Helmut Buske.

Jakobi, Angelika & Ali Ibrahim. 2018. Labile verbs in Tagle. In Gerrit J. Dimmendaal, Birgit Hellwig & Gertrud Schneider-Blum (eds.), *Nuba Mountain Language Studies: New Insights,* 99–127. Cologne: Rüdiger Köppe.

Krell, Amy. 2012. Rapid Appraisal Sociolinguistic Survey Among Ama, Karko, and Wali Language Groups. *SIL Electronic Survey Report 2012-023*. Dallas, Texas: SIL International.

Kutsch Lojenga, Constance. 1994. *Ngiti: A Central-Sudanic Language of Zaire*. Cologne: Rüdiger Köppe.

Ladd, D. Robert, Bert Remijsen & Caguor Adong Manyang. 2009. On the distinction between regular and irregular inflectional morphology: Evidence from Dinka. *Language* 85. 659–670.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2015. *Ethnologue, Languages of Africa and Europe*, 18th edn. Dallas, Texas: SIL International.

Moodie, Jonathan. 2019. A Grammar of Lopit. PhD thesis, University of Melbourne.

Stevenson, Roland C. 1984. *The Nuba People of Kordofan Province. An Ethnographic Survey* (Graduate College Publications, Monograph 7). Khartoum: University of Khartoum.

Storch, Anne. 2005. *The Noun Morphology of Western Nilotic*. Cologne: Rüdiger Köppe.

Tucker, A. N. & M. A. Bryan. 1966. *Linguistic Analyses. The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press for International African Institute.

Waag, Christine. 2010. *The Fur Verb and its Context*. Cologne: Rüdiger Köppe.

Weiss, Doris. 2009. *Phonologie et morphosyntaxe du maba*. Thèse de doctorat, Université Lumière Lyon 2.

Denis Creissels

4 Number in Tswana

Abstract: This chapter describes various aspects of the number category in Tswana, a Southern Bantu language spoken in Botswana and South Africa. Number is an obligatory category of Tswana common nouns. Number marking is morphologically complex and closely related to the gender system. Gender-number agreement is pervasive in Tswana morphosyntax. The semantic organization of the number category in Tswana is broadly similar to that found in Standard Average European languages.

1 Overview

1.1 The language

Tswana (aka Setswana), S31 in Guthrie's nomenclature of Bantu languages, is a Southern Bantu language with more than 6 million speakers. In Botswana, ethnic Batswana constitute 80 % of the population, estimated at 2.3 million (2020). In South Africa, Tswana is dominant in the Northwest Province and in some districts of the Free State Province, and the number of its speakers is estimated at 4–5 million. The closest relatives of Tswana are Pedi and Southern Sotho. Tswana, Pedi and Southern Sotho are so close to each other that, from a strictly linguistic point of view, they may be considered as three varieties of a single language. Pedi is commonly designated as Northern Sotho, but this term is ambiguous, since it is also used with reference to some lects (Lobedu, Tswapong, etc.) that, linguistically, are better considered languages distinct from Sotho-Tswana proper.

Typologically, Tswana is in almost every respect a typical Bantu language.[1] Its locative system, very different from that found in Central Bantu languages, is the main exception to this generalization, but this is an areal feature, shared by the other Southern Bantu languages.[2]

1.2 The basics of Tswana morphosyntax

1.2.1 Nouns and noun phrases

Definiteness marking is not grammaticalized in Tswana. There is no distinction between a definite and an indefinite form of nouns, and NPs including no determiner

---

1 On the typological profile of Bantu languages, cf. Creissels (Forthcoming).

2 For an overall presentation of Tswana, see Cole (1955), Creissels (2003). On Southern Sotho, see Doke & Mofokeng (1957). On the changes in the organization of the locative system that affected Southern Bantu languages, see Marten (2010), Creissels (2011).

can be used freely to introduce new referents or to represent discursively salient referents, or referents that can be viewed as identifiable by the addressee in one way or another.

In Tswana, nominal classification does not take the form of East Asian-style classifiers, but of a multiple-gender system of the type commonly found in Bantu (and more generally Niger-Congo) languages, in which the masculine vs. feminine distinction plays absolutely no role. Tswana has nine distinct genders. Two major characteristics of this type of gender system that are directly relevant to the topic of this article are that (a) the division of nouns into genders is closely correlated to the division of nouns into inflectional classes according to the pairs of obligatory prefixes marking the singular vs. plural distinction, and (b) gender agreement markers and number agreement markers are completely fused and show no evidence of ever having been dissociated, hence the traditional notion of 'class agreement' conflating gender and number agreement.

In this chapter, Tswana genders are designated according to the convention generally followed in Bantu studies, according to which numbering refers to the agreement classes of noun forms reconstructed for Proto-Bantu. An obvious advantage of using arbitrary numbers as labels for genders is that the use of semantically motivated labels would be hard to reconcile with the semantic heterogeneity of genders. On the content of Tswana genders, cf. Cole (1955: 68–101).

In Tswana, noun *forms* divide into 12 agreement classes, and each of the 9 genders can be defined as a set of nominal *lexemes* that have the same agreement behavior both in the singular and the plural. For example, the singular form *mosadi* **mù-sádí** 'woman' belongs to the agreement class of noun forms traditionally labeled class 1, the corresponding plural *basadi* **bà-sádí** 'women' belongs to an agreement class of noun forms traditionally designated as class 2, and consequently *mosadi* **mùsádí** *as a lexeme* belongs to a gender that can be designated as gender 1–2.[3] *Monna* **mù-ńná** 'man' pl. *banna* **bà-ńná**, or *ntšhe* **ǹtʃʰé** 'ostrich' pl. *bontšhe* **bó-ǹtʃʰé**, have exactly the same agreement properties as *mosadi* **mùsádí**, both in the singular and in the plural, and consequently also belong to gender 1–2. The singular form *motse* **mù-tsì** 'village' shows the same singular prefix as *mosadi* **mù-sádí** 'woman' or *monna* **mù-ńná** 'man', but belongs to a distinct agreement class of noun forms, traditionally labeled class 3, whereas the corresponding plural form *metse* **mì-tsì** belongs to an agreement class traditionally labeled class 4, and consequently *motse* **mùtsì** *as a lexeme* belongs to gender 3–4. *Legodu* **lì-χòdù** 'thief' pl. *magodu* **mà-**

---

3 In this article, Tswana forms are systematically quoted both according to current Tswana orthography (in italics), and in broad phonetic transcription (in bold). The reason is that current orthography may be quite misleading in a linguistic analysis, since it distinguishes only 5 vowels and does not indicate tones at all, whereas Tswana has 9 vowel phonemes, and tones are crucial for morphological analyses. Moreover, many morphemes that are unquestionably prefixes (in particular, subject indexes and object indexes) are written as if they were separate words. In the examples, the correct word division is given in the phonetic transcription (second line).

χòdù belongs to another gender, traditionally labelled 5–6, whose singular forms belong to agreement class 5, whereas the corresponding plural forms belong to agreement class 6.[4]

As illustrated by example (1), in which *mosadi* **mʊ̀-sádí** 'woman' (gender 1–2) and *lekau* **lì-káú** 'boy' (gender 5–6) combine with two adjectives, a relative clause, and a demonstrative, Tswana NPs have two very general characteristics: noun dependents follow their head, and express gender and number agreement with their head.

(1) a. *mosadi yo moleele yo montsho yo o opelang yole*
    mʊ̀-sádì   **jó**   **mʊ̀-léèlé** **ꜝjó**   **mʊ́-ǹtsʰʊ̀** **jó**
    SG-woman (1)[5] CL1.ATTR CL1-tall   CL1.ATTR CL1-black CL1.ATTR
    ꜝʊ́-ɔ́pɛ́là-ŋ́   **ꜝjó-lé**
    S:CL1-sing:PRS-REL DEM.CL1-DIST
    'that tall woman with dark complexion who is singing'

    b. *basadi ba baleele ba bantsho ba ba opelang bale*
    bà-sádì   **bá**   **bà-léèlé** **ꜝbá**   **bá-ǹtsʰʊ̀** **bá**
    PL-woman(2) CL2.ATTR CL2-tall CL2.ATTR CL2-black CL2.ATTR
    ꜝbá-ɔ́pɛ́là-ŋ́   **ꜝbá-lé**
    S:CL2-sing:PRS-REL DEM.CL2-DIST

    c. *lekau le leleele le lentsho le le opelang lele*
    lì-káù   **lé**   **lì-léèlé** **ꜝlé**   **lí-ǹtsʰʊ̀** **lé**
    SG-boy(5) CL5.ATTR CL5-tall CL5.ATTR CL5-black CL5.ATTR
    ꜝlí-ɔ́pɛ́là-ŋ́   **ꜝlé-lé**
    S:CL5-sing:PRS-REL DEM.CL5-DIST

---

4 Traditionally, in Bantu linguistics, the choice of a particular number to designate a given agreement pattern is determined by the relationship between the gender-number agreement patterns of present-day Bantu languages and the gender-number agreement patterns reconstructed for Proto-Bantu. Note in particular that gaps in the numbering of the agreement patterns of Tswana are due to the fact that Tswana has no direct reflex of some of the agreement patterns reconstructed for Proto-Bantu. Note also that, for the sake of simplicity, in this article, I designate simply as 'class 10' and 'class 17' two agreement patterns that could be designated as 'class 8/10' and 'class 15/17' to better account for their relationships with reconstructed Proto-Bantu classes. On the reconstruction of Proto-Bantu inflectional classes of nouns and agreement patterns, cf. Meeussen (1967).

5 Contrary to the tradition in Bantu studies, the gloss CL ('noun class') is used here for gender-number agreement markers only. Noun prefixes are glossed SG (singular) or PL (plural), and the agreement pattern governed by the noun form is indicated by a number between parentheses after the lexical gloss. The number that follows CL in the glosses of gender-number agreement markers also refers to one of the 12 possible agreement patterns that may be triggered by Tswana noun forms.

d. *makau a maleele a mantsho a a opelang ale*

mà-káù   á   mà-léèlé ꜜá   má-ǹtsʰʊ̀ á
PL-boy(6) CL6.ATTR CL6-tall   CL6.ATTR CL6-black CL6.ATTR
ꜜá-ɔ́pɛ́là-ŋ́   ꜜá-lé
S:CL6-sing:PRS-REL DEM.CL6-DIST

There is no exception to the rule of obligatory gender-number agreement between noun dependents and their head. In contrast, the *head-dependent* order is not absolutely obligatory, although noun dependents preceding their head are extremely rare in spontaneous texts. The anteposition of noun dependents adds emphasis, but never modifies the denotative meaning. Anteposition is equally possible (and equally rare) with all types of noun dependents.

Noun dependents divide into subtypes according to the particular sets of agreement markers by means of which they express gender-number agreement.

1.2.2 Verbs

Verb inflection involves TAM markers, negation markers, subject indexes, object indexes, and conjoint/disjoint markers.[6] Tswana verbs are also characterized by a rich system of valency-changing derivations (reflexive, reciprocal, decausative, causative, passive, and applicative) – see Creissels (2002, 2006). The tonal morphology of the verb is particularly complex, even by Bantu standards – cf. Creissels (1999), Creissels & al. (1997), Creissels (2017).

1.2.3 Clauses

Basic verbal clauses have a rigid constituent order *subject – verb – objects – obliques*. Alignment is consistently nominative/accusative: the coding frames through which verbs express their argument structure invariably include a term whose coding properties are identical to those of the agent of prototypical transitive verbs. There is no case contrast, but the distinction between subjects and objects is apparent in indexation: with only the exception of the infinitive and the imperative, verb forms include

---

6 In the context of Bantu studies, a conjoint verb form is a verb form that cannot be found in sentence-final position and cannot be separated from the following phrase by a pause. A disjoint verb form does not have this limitation, but is not excluded from non-final contexts either, and when in non-final sentence position, is not necessarily followed by a perceptible pause. There is some cross-linguistic variation in the function of the conjoint/disjoint contrast – Van der Wal & Hyman (2017). In Tswana, it marks the distinction between phrases in post-verbal position that form part of the verb phrase and contribute to the comment expressed by the verb, and phrases in post-verbal position that fulfill the discourse function of afterthought (alias antitopic) – Creissels (2017).

an obligatory subject index, even in the presence of a subject noun phrase; objects can be indexed too, but object indexes occupy a distinct position in the verbal template, and occur only if the clause includes no noun phrase representing the same argument.

As illustrated in (2), if a co-referent NP is present, the subject index expresses person-gender-number agreement with it. The subject NP is syntactically optional, and in its absence, subject indexes that do not belong to 1st or 2nd person are interpreted anaphorically, triggering the identification of the argument they represent to a contextually salient referent compatible with the gender-number value expressed by the subject index.

(2) a. *Ngwana o thubile mae.*
    ŋw-àná ú-tʰùbílé mà:-í.
    SG-child(1) S:CL1-break:PRF:CJ PL-egg(6)
    'The child broke the eggs.'

    b. *Ngwana o tsile.*
    ŋw-àná ꜜú-tsî:lè.
    SG-child(1) S:CL1-come:PRF:DJ
    'The child came.'

    c. *O thubile mae.*
    ú-tʰùbílé mà:-í.
    S:CL1-break:PRF:CJ PL-egg(6)
    'He/She broke the eggs.'

    d. *O tsile.*
    ú-tsî:lè.
    S:CL1-come:PRF:DJ
    'He/she came.'

    e. **Ngwana thubile mae.*

    f. **Ngwana tsile.*

As illustrated in (2a) and (2c), the object of transitive verbs is not obligatorily indexed on the verb form, but topical objects whose precise description is considered superfluous by the speaker are represented by object indexes prefixed to verbs. Object indexes immediately precede the verb stem and may be separated from subject indexes by TAM or negation markers, as in (3a–b).

(3) a. *Ngwana o a thubile.*
    ŋw-àná ú-à-tʰúbî:lè.
    SG-child(1) S:CL1-O:CL6-break:PRF:DJ
    'The child broke them (the eggs).'

b. *Ngwana o tlaa a thuba.*
ŋw-àná **ꜝʊ́-tɬáà-á-tʰû:bà.**
SG-child(1) S:CL1-FUT-O:CL6-break:DJ
'The child will break them (the eggs).'

Tswana has very productive multiple-object constructions: non-derived verbs may have two non-coordinated objects expressing distinct semantic roles, and valency-increasing derivations (causative and applicative) may result in constructions with three objects – example (4). In multiple-object constructions, the syntactic differences between the objects are minimal: each object can be converted into the subject of a passive construction, or represented by an object index. The linear order of the objects is rigid, and depends crucially on Animacy Hierarchy.

(4) a. *Ngwana o nole maši.*
ŋw-àná **ꜝʊ́-nʊ́lé** **mâ:-ʃì.**
SG-child(1) S:CL1-drink:PRF:CJ PL-milk(6)
'The child drank milk.'

b. *Ke nositse ngwana maši.*
kì-nʊ́sítsé **ŋw-àná** **mâ:-ʃì.**
S:1SG-drink:CAUS:PRF:CJ SG-child(1) PL-milk(6)
'I made the child drink milk.'

c. *Ke noseditse Dimpho ngwana maši.*
kì-nʊ́sédítsé **dím̀pʰɔ́** **ŋw-àná** **mâ:-ʃì.**
S:1SG-drink:CAUS:APPL:PRF:CJ Dimpho(1) SG-child(1) PL-milk(6)
'I made the child drink milk in Dimpho's place.'

Obliques are commonly expressed as prepositional phrases, but due to (a) a very high degree of transitivity prominence and (b) the productivity of valency-increasing derivations and multiple-object constructions, many participants encoded as obliques in Standard Average European languages are encoded in Tswana as object NPs. On Tswana prepositions, cf. Creissels (2013)

1.2.4 Overview of the number category in Tswana

If inflection is defined as the part of morphology that may be directly involved in the application of syntactic rules, the inflectional nature of number in Tswana follows from the fact that agreement in number (and gender) is pervasive in Tswana NPs and clauses. With no exception, noun modifiers agree in number and gender with their head, and verbs agree in person and number (and in the third person, in gender) with their subject.

As already mentioned in 1.2.1, Tswana has a gender system characterized by a relatively high number of genders (9) and the kind of interaction with number marking commonly found in the Niger-Congo languages whose gender system is traditionally described in terms of 'noun classes'.

Number is obligatory in the sense that the overwhelming majority of Tswana nouns have a singular form and a plural form, and in most contexts, singular forms are obligatorily interpreted as referring to a single individual. On this point, the situation in Tswana is basically the same as in Standard Average European languages. For example, the clauses in (5) provide no indication about the identifiability of the referent of *kgomo* **qʰòmʊ́** 'cow' pl. *dikgomo* **dì-qʰòmʊ́**, but (5a) implies reference to a situation in which no more than one cow is present, whereas (5b) implies reference to a plurality of cows.

(5) a. *Ke bonye kgomo ko nokeng.*
     kì-bóɲí   qʰòmʊ́   kó nòkê:-ǹ.
     S:1SG-see:PRF:CJ (SG)cow(9) at (SG)river(9)-LOC
     'I saw a/the cow at the river.'

    b. *Ke bonye dikgomo ko nokeng.*
     kì-bóɲí   dí-qʰòmʊ́   kó nòkê:-ǹ.
     S:1SG-see:PRF:CJ PL-cow(10) at (SG)river(9)-LOC
     'I saw (the) cows at the river.'

Number marking is complex. Each of the nine genders is characterized by a particular pair of prefixes acting as singular and plural markers, there is no one-to-one correspondence between singular and plural markers, and number agreement cannot be dissociated from gender agreement.

Number is a binary category (singular vs. plural). Notions such as 'dual' or 'paucal' are not grammaticalized in Tswana.

It is possible to use gender shift to express a collective meaning, but this mechanism is not very productive.

Tswana has an associative plural marker **bó-** that can combine with proper nouns and kinship terms. The same marker **bó-** is also used as an ordinary plural marker with some common nouns that have a zero prefix in the singular (see (9) below).

Verbal number does not exist in Tswana.

2 Pronominal, nominal, and verbal number

2.1 Generalities

The only asymmetry between nouns and pronouns concerns plural exponence in first and second person pronouns (cf. 2.2). Tswana has obligatory subject-verb agreement, but does not have verbal number.

2.2 Pronominal number

First and second person pronouns can be segmented into a person-number prefix and a stem **-ná**. As can be seen in (6), it is not possible to segment the prefix into a person marker and a plural marker.

(6) first and second person pronouns

	sg.	pl.
1st	*nna*	*rona*
	ǹ-ná	**rʊ̀-ná**
2nd	*wena*	*lona*
	wɛ̀-ná	**lʊ̀-ná**

In the third person, pronouns agree in gender and number with their antecedent. With the exception of the singular of gender 1–2, they can be segmented into a stem **-ɔ̀nɛ́** and a gender-number prefix. Note that the nouns in gender 17 do not vary in number, which explains the existence of a single form of the third person pronoun for this gender

(7) third person pronouns

	sg.	pl.
gender 1–2	*ene*	*bone*
	ɛ̀nɛ́	**b-ɔ̀nɛ́**
gender 3–4	*one*	*yone*
	Ø-ɔ̀nɛ́	**j-ɔ̀nɛ́**
gender 5–6	*lone*	*one*
	l-ɔ̀nɛ́	**Ø-ɔ̀nɛ́**
gender 7–10[7]	*sone*	*tsone*
	s-ɔ̀nɛ́	**ts-ɔ̀nɛ́**

---

7 Tswana has lost the distinction between the Proto-Bantu classes 8 and 10. This means that the plural prefix ***bi** of the nouns whose plural historically belonged to class 8 has been replaced in Tswana by the plural prefix of class 10, and their agreement pattern has changed accordingly.

gender 9–10	*yone*	*tsone*
	j-ɔ̀nɛ́	**ts-ɔ̀nɛ́**
gender 11–6	*lone*	*one*
	l-ɔ̀nɛ́	**Ø-ɔ̀nɛ́**
gender 11–10	*lone*	*tsone*
	l-ɔ̀nɛ́	**ts-ɔ̀nɛ́**
gender 14–6	*jone*	*one*
	dʒ-ɔ̀nɛ́	**Ø-ɔ̀nɛ́**
gender 17[8]		*gone*
		χ-ɔ̀nɛ́

The interrogative pronouns *mang* **má́ŋ** 'who?' and *eng* **ìŋ** 'what?' are invariably used pronominally. *Mang* **máŋ** 'who?' has a plural form *bomang* **bó-máŋ**, whereas *eng* **ìŋ** 'what?' is invariable. The other types of pronouns are words also used adnominally. Their gender-number inflection expresses the same distinctions as that of third person pronouns. For example, *-pe* **-pɛ́** 'any' has the inflection shown in (8), both in its adnominal and pronominal use.

(8) gender-number inflection of *-pe* **-pɛ́** 'any'

	sg.	pl.
gender 1–2	*ope*	*bape*
	ʊ́-pɛ́	**bá-pɛ́**
gender 3–4	*ope*	*epe*
	ʊ́-pɛ́	**í-pɛ́**
gender 5–6	*lepe*	*ape*
	lí-pɛ́	**á-pɛ́**
gender 7–10	*sepe*	*dipe*
	sí-pɛ́	**dí-pɛ́**
gender 9–10	*epe*	*dipe*
	í-pɛ́	**dí-pɛ́**
gender 11–6	*lope*	*ape*
	lʊ́-pɛ́	**á-pɛ́**
gender 11–10	*lope*	*dipe*
	lʊ́-pɛ́	**dí-pɛ́**

8 In Tswana, this gender includes the infinitives and the two nouns *golo* **χʊ̀lɔ̀** 'place' and *felo* **fìlɔ̀** place'. The nouns whose singular historically belonged to class 15 have been transferred to other genders, and the same occurred with most of the locative nouns that historically belonged to the locative classes (16, 17, and 18). Moreover, the distinction between the locative agreement patterns 16, 17 and 18 has been lost in Tswana, as in other Southern Bantu languages. For more details on the changes that have affected the system of locative classes in Tswana, se Creissels (2011).

gender 14–6	*bope*	*ape*
	búpɛ́	**á-pɛ́**
gender 17		*gope*
		χʊ́-pɛ́

In Tswana, as mentioned above, subjects are obligatorily indexed, whereas the indexation of objects is syntactically optional. Subject and object indexes vary in person, number and gender exactly like personal pronouns (cf. 3.4).

2.3 Nominal number

2.3.1 Number inflection of nouns

The Tswana nouns that have a singular vs. plural distinction divide into 9 inflectional classes depending on the pair of prefixes they select as singular and plural markers. The following table gives the correspondence between pairs of number markers and genders.[9] Gender 1–2 conflates two groups of nouns selecting distinct number prefixes, whereas the correspondence between the other genders and number inflection is a straightforward one-to-one correspondence. As can be seen in (9), the labeling of genders by means of numbers referring to agreement classes of noun forms makes immediately apparent the neutralization of some gender distinctions either in the singular (singular forms belonging to agreement class 11 may correspond to plural forms of class 6 or 10) or in the plural (plural forms belonging to agreement class 6 may correspond to singular forms of class 5, 11, or 14, and plural forms belonging to agreement class 10 may correspond to singular forms of class 7, 9, or 11).

(9) number inflection of nouns and genders

number markers	gender	examples
mʊ̀- (sg.) / **bà-** (pl.)	1–2	*mosadi* **mʊ̀-sádí** 'woman' pl. *basadi* **bà-sádí**
Ø- (sg.) / **bó-** (pl.)	1–2	*ntšhe* **ɲ̀tʃʰé** 'ostrich' pl. *bontšhe* **bóɲ̀tʃʰé**
mʊ̀- (sg.) / **mì-** (pl.)	3–4	*motse* **mʊ̀-tsì** 'village' pl. *metse* **mì-tsì**
lì- (sg.) / **mà-** (pl.)	5–6	*lee* **lì-í** 'egg' pl. *mae* **mà-í**
sì- (sg.) / **dì-** (pl.)	7–10	*selepe* **sì-lɛ́pɛ́** 'axe' pl. *dilepe* **dì-lɛ́pɛ́**
Ø- (sg.) / **dì-** (pl.)	9–10	*podi* **pʊ́dí** 'goat' pl. *dipodi* **dì-pʊ́dí**

---

9 A pattern similar to that described here for Tswana is found in most other Bantu languages (for example, Swahili). There is however variation due to the fact that pairs of classes reconstructed as distinct classes in Proto-Bantu (for example, class 3 and class 14, or class 8 and class 10) may have variously merged in the individual languages.

> **lʊ̀-** (sg.) / **mà-** (pl.) 11–6 *losea* **lʊ̀-síá** 'baby' pl. *masea* **mà-síá**
> **lʊ̀-** (sg.) / **dì-** (pl.) 11–10 *loso* **lʊ̀-sɔ̀** 'spoon' pl. *dintsho* **dì-ǹtsʰɔ̀**
> **bʊ̀-** (sg.) / **mà-** (pl.) 14–6 *bothata* **bʊ̀-tʰátá** 'problem' pl. *mathata* **mà-tʰátá**

As already mentioned above, the nouns belonging to gender 17, characterized by an agreement pattern that coincides with none of the agreement patterns found with the singular or plural forms of any other gender, do not vary in number. This gender includes the infinitives (formed by adding the prefix **χʊ̀-** to verb stems), plus the two nouns *golo* **χʊ̀-lɔ̂** 'place' and *felo* **fʊlɔ̂** 'place'.[10]

2.3.2 Singularia tantum, pluralia tantum

Much in the same way as in Standard Average European languages, in Tswana, number inflection is not limited to nouns relatively high on animacy hierarchy, and nouns lending themselves to number inflection can be found among all semantic types of nouns. Animacy hierarchy is relevant in the sense that the nouns that stand high in the hierarchy are more prone to be treated as count nouns (with obligatory plural marking when referring to more than one individual) than those standing lower, but it is not possible to formulate strict rules.

There are however nouns that do not have the singular vs. plural distinction. The case of the nouns belonging to gender 15–17 has already been evoked in 2.3.1, but among the nouns that do not lend themselves to the singular vs. plural distinction, there are also nouns with prefixes and agreement properties that coincide with those of either the singular or the plural of one of the other genders. I am aware of no in-depth study of this question on the basis of which precise generalizations could be put forward. However, as can be seen from the following examples, this behavior seems to be typically found with nouns for which the singular vs. plural distinction is more or less problematic conceptually:[11]

- *lenyora* **lì-ɲɔ́rá** 'thirst' has the same prefix and agreement properties as singular forms of gender 5–6,
- *mala* **mà-lá** 'cold' and *madi* **mà-dí** 'blood' have the same prefix and agreement properties as plural forms of gender 5–6,
- *serame* **sì-ràmɛ́** 'cold' has the same prefix and agreement properties as singular forms of gender 7–10,

---

10 *Felo* **fʊlɔ̂** 'place' was originally formed by adding the prefix of Bantu class 16 to a stem cognate with that of its synonym *golo* **χʊ̀-lɔ̂**, but synchronically, due to the obsolescence of the Bantu class 16 in Tswana, *felo* **fʊlɔ̂** cannot be segmented anymore, and has been re-assigned to the same agreement class as *golo* **χʊ̀-lɔ̂**. This irregularity tends to be eliminated by the replacement of *felo* **fʊlɔ̂** by a regular noun of gender 5–6: *lefelo* **lì-fʊ̀lɔ̂** pl. *mafelo* **mà-fʊlɔ̂**.
11 On the names of groups, see §2.3.5.

- *bosadi* **bʊ̀-sádí** 'womanhood' (abstract noun derived from *mosadi* **mʊ̀-sádí** 'woman' by gender shift)[12] and *botlhabatsatsi* **bʊ̀-tɬʰàbà-tsàtsí** 'east' (cf. *tlhaba* **tɬʰàbà** 'pierce' and *letsatsi* **lì-tsàtsí** 'sun') have the same prefix and agreement properties as singular forms of gender 14–6,
- etc.

2.3.3 Special readings of number inflection

The nouns whose default behavior is that of mass nouns, for which the singular vs. plural distinction is not relevant, lend themselves to the cross-linguistically common types of 'special readings' of number inflection. For example:
- *bojalwa* **bʊ̀-dʒàlwá** 'traditional beer' is most commonly used in the singular as a mass noun, but the plural form *majalwa* **mà-dʒàlwá** is available to express 'different kinds of beer' or 'several sets of beer containers',
- 'food' is normally expressed by the plural form *dijo* **dì-dʒɔ́**, but the corresponding singular form *sejo* **sì-dʒɔ́** can be used with the meaning 'a particular kind of food',
- 'money' is expressed by the plural form *madi* **mà-dí** (homonymous with *madi* **mà-dí** 'blood'), but the corresponding singular form *ledi* **lì-dí** can be used with the meaning 'coin',
- etc.

In generic clauses such as 'Bears hibernate', nouns are most commonly in the plural form, i.e. in the same form as when they are used to refer to specific groups of individuals.

(10) *Ditau di tshela ka go tsoma diphologolo tse dingwe.*
    dì-tàú   dí-tsʰílà   ká   χʊ̀-tsʊ́má dì-pʰɔ́lɔ̀χɔ́lɔ̀ tsé
    PL-lion(10) S:CL10-live:PRS:CJ with INF-hunt PL-animal(10) CL10.ATTR
    dìː-ŋwí.
    CL10-other
    'Lions make their living by hunting other animals.'

2.3.4 Associative plural

The prefix **bó-** is the regular plural marker for the nouns of gender 1–2 that have a zero prefix in the singular, such as *ntšhe* **ɲ̀tʃʰé** 'ostrich' pl. *bontšhe* **bó-ɲ̀tʃʰé**, or

---

12 The case of *bosadi* **bʊ̀-sádí** 'womanhood' is illustrative of abstract nouns in general: their formation involves no specific derivational morphology, just gender shift by which nouns of individuals belonging to any other gender are converted into singularia tantum of gender 14–6 expressing

kgengwe **qʰɛ́ŋwɛ́** 'tsamma melon' pl. *bokgengwe* **bó-qʰɛ́ŋwɛ́**, but the same prefix **bó-** can be used as an associative plural marker with proper names and kinship terms.

With proper names, associative plural is the only possible interpretation of **bó-** (as in *boMpho* **bó-m̀pʰɔ́** 'Mpho and others'), whereas with kinship terms, there is ambiguity between the associative plural and ordinary plural readings, since kinship terms belong to the subset of gender 1–2 nouns that have a zero prefix in the singular, cf. for example *malome* **màlʊ́mɛ́** 'my maternal uncle' / *bomalome* **bó-màlʊ́mɛ́** 'my maternal uncles' or 'my maternal uncle and others'.

Associative plural marking is involved in a cross-linguistically rare type of inclusory coordination. Tswana does not have inclusory coordination with a plural pronoun as the first term of the construction. In the Tswana inclusory coordination construction, the first term of the construction is the associative plural of a noun. For example, the literal meaning of *boMpho le Kitso* **bó-m̀pʰɔ́ ˈlí-kítsɔ́** is 'Mpho and others with Kitso', but this construction is commonly used as referring to a group consisting of two individuals only, Mpho and Kitso. For more details on the prefix used to express associative plural and on the inclusory coordination construction, cf. Creissels (2016).

2.3.5 Group nouns (collectives)

Nouns with intrinsically plural reference can be found in various genders, e.g. *motlhape* **mʊ̀-tɬʰàpí** 'herd' pl. *metlhape* **mì-tɬʰàpí** (gender 3–4), *morafe* **mʊ̀-ráfí** 'tribe' pl. *merafe* **mì-ráfí** (gender 3–4) or *letsomane* **lì-tsʊ́mání** 'flock of sheep or goats' pl. *matsomane* **mà-tsʊ́mání** (gender 5–6). They form plurals and agree exactly like the other nouns showing the same inflectional prefixes.

Group nouns can be productively formed from nouns referring to animate individuals by mere gender shift accompanied by the corresponding change in the inflectional prefixes, without the use of any overt derivational marker. Group nouns referring to humans are found in gender 9–10 (for example *nna* **Ø-ńná** 'crowd of men' < *monna* **mʊ̀-ńná** 'man'), whereas group nouns referring to animals are found in gender 11–6 (for example *lobotsane* **lʊ̀-bʊ́tsàní** 'flock of goat-kids' < *potsane* **pʊ́tsàní** 'goat-kid', or *lolau* **lʊ̀-làú** 'pride of lions' < *tau* **tàú** 'lion', with a stem-initial alternation triggered by the change of prefix).[13]

---

abstract quality, as for example *tau* **tàú** 'lion' pl. *ditau* **dì-tàú** (gender 9–10) > *bolau* **bʊ̀-làú** 'leonine nature' (class 14), where the consonant alternation **t ~ l** is the result of an automatic morphophonological process operating at the junction between noun stems and prefixes.

13 The explanation of this alternation is that, historically, nouns had a prefix n- in gender 9–10, as opposed to the CV prefixes found in the other genders.

In the plural, group nouns derived via gender shift may have a 'greater plural' reading, for example *malau* **mà-làú** (plural of *lolau* **lʊ̀-làú** 'pride of lions') 'very many lions' (Cole 1955: 100).

2.3.6 Number marking of nominalized adnominal modifiers

When nominalized, adnominal modifiers show exactly the same gender-number marking as when they agree with their head in a noun-modifier construction.

2.4 Verbal number

Verbal number as a category encompassing repetition of action and plurality of participants does not exist in Tswana. Iterative verbs can be formed via reduplication (as for example *ragaraga* **ráχáráχà** 'kick repeatedly' < *raga* **ráχá** 'kick', but verb reduplication is not used to encode plurality of participants.

3 Agreement and the syntax of number

3.1 Number agreement in noun-modifier constructions

In Tswana, all noun-modifier constructions without any exception are subject to agreement in gender and number. Depending on the nature of the modifier, the agreement marks may appear on the modifier itself (for example, with demonstratives, or the interrogative determiner *-fe* **-fí** 'which'), on a linker introducing the modifier (for example, with adnominal possessors, or non-agreeing numerals), or both (for example, with adjectives, or agreeing numerals). (11) illustrates the agreement of the interrogative determiner, and (12) illustrates the agreement of attributive adjectives and of the attributive linker obligatorily inserted between attributive adjectives and their head.

(11)   The agreement of the interrogative determiner

gender	singular	plural
1–2	*mosadi ofe*	*basadi bafe*
	mʊ̀-sádì ʊ́-fí	**bà-sádì bá-fí**
	'which woman?'	'which women?'
	ntšhe ofe	*bontšhe bafe*
	ɲ̀tʃʰé ʊ́-fí	**bó-ɲ̀tʃʰé bá-fí**
	'which ostrich?'	'which ostriches?'

3–4	*molemo ofe*	*melemo efe*
	mʊ̀-lìmɔ̀ ʊ́-fí	**mì-lìmɔ̀ í-fí**
	'which medecine?'	'which medecines?'
5–6	*legodu lefe*	*magodu afe*
	lì-χòdù lí-fí	**mà-χòdù á-fí**
	'which thief?'	'which thieves?'
7–10	*sekolo sefe*	*dikolo dife*
	sì-kólò sí-fí	**dì-kólò dí-fí**
	'which school?'	'which schools?'
9–10	*kgosi efe*	*dikgosi dife*
	qʰósì í-fí	**dì-qʰósì dí-fí**
	'which chief?'	'which chiefs?'
11–6	*losea lofe*	*masea afe*
	lʊ̀-síà lʊ́-fí	**mà-síà á-fí**
	'which baby?'	'which babies?'
11–10	*lokwalo lofe*	*dikwalo dife*
	lʊ̀-kwálɔ̀ lʊ́-fí	**dì-kwálɔ̀ dí-fí**
	'which book?'	'which books?'
14–6	*botshelo bofe*	*matshelo afe*
	bʊ̀-tsʰílɔ̀ bʊ́-fí	**mà-tsʰílɔ̀ á-fí**
	'which life?'	'which lives?'
17		*golo gofe*
		χʊ̀-lɔ̀ χʊ́-fí
		'which place?'

(12) The agreement of attributive adjectives and of the attributive linker

gender	singular	plural
1–2	*mosadi yo moša*	*basadi ba baša*
	mʊ̀-sádì jó mʊ̀-ʃá	**bà-sádì bá bà-ʃá**
	'new woman'	'new women'
	ntšhe yo moša	*bontšhe ba baša*
	ɲ̀ʧʰé jó mʊ̀-ʃá	**bó-ɲ̀ʧʰé bá bà-ʃá**
	'new ostrich'	'new ostriches'
3–4	*molemo o moša*	*melemo e meša*
	mʊ̀-lìmɔ̀ ó mʊ̀-ʃá	**mì-lìmɔ̀ é mì-ʃá**
	'new medecine'	'new medecines'
5–6	*legodu le leša*	*magodu a maša*
	lì-χòdù lé lì-ʃá	**mà-χòdù á mà-ʃá**
	'new thief'	'new thieves'
7–10	*sekolo se seša*	*dikolo tse dintšha*
	sì-kólò sé sì-ʃá	**dì-kólò tsé dí-ɲ̀ʧʰá**
	'new school'	'new schools'

9–10	*kgosi e ntšha*	*dikgosi tse dintšha*
	qʰósì é ɲ̀ʧʰá	**dì-qʰósì tsé dí-ɲ̀ʧʰá**
	'new chief'	'new chiefs'
11–6	*losea lo loša*	*masea a maša*
	lʊ̀-síà ló lʊ̀-ʃá	**mà-síà á mà-ʃá**
	'new baby'	'new babies'
11–10	*lokwalo lo loša*	*dikwalo tse dintšha*
	lʊ̀-kwálɔ̀ ló lʊ̀-ʃá	**dì-kwálɔ̀ tsé dí-ɲ̀ʧʰá**
	'new book'	'new books'
14–6	*botshelo jo boša*	*matshelo a maša*
	bʊ̀-tsʰîlɔ̀ ʤó bʊ̀-ʃá	**mà-tsʰîlɔ̀ á mà-ʃá**
	'new life'	'new lives'
17		*golo mo goša*
		χʊ̀-lɔ̀ mó χʊ̀- ʃá
		'new place'

3.2 Numeral modification and number

In Tswana, when modified by numerals other than 'one' or by other quantifiers denoting plurality, nouns are obligatorily in the plural. Some numeral modifiers agree in gender and number with their head, others don't, but in all cases, they are obligatorily introduced by a linker expressing gender-number agreement.

Measure words combine with numbers like ordinary nouns. For example, in (13), *litara* **lítàrá** 'liter' pl. *dilitara* **dì-lítàrá** behaves exactly like any other noun of gender 9–10 modified by a numeral and an adnominal possessor.

(13) a. *litara ya maši*
  lítàrá   **ˈy-á-má-ʃì**
  (SG)liter(9) CL9-GEN-PL-milk(6)
  'a liter of milk'

  b. *dilitara di le pedi ya maši*
  dì-lítàrá   **ˈdí-lí**   **pèdí**   **ˈts-á-má-ʃì**
  PL-liter(10) CL10-NUM (CL10)two CL10-GEN-PL-milk(6)
  'two liters of milk'

3.3 NP coordination and number

In Tswana, the additive coordination of NPs is expressed as *NP₁ le NP₂*, where *le* **lí-** 'and' is a proclitic also found in other constructions with the meanings 'with' and 'even'.

As can be expected, if one of the coordinands is a 1st person pronoun, *NP₁ le NP₂* triggers 1st person plural agreement, irrespective of the nature of the second coordinand.[14] If one of the coordinands is a 2nd person pronoun, and the other is not a 1st person pronoun, *NP₁ le NP₂* invariably triggers 2nd person plural agreement.

(14) a. *Nna le Kitso re tlaa sala mono.*
      ǹná ˈlí-kítsɔ́ rí-tɬàà-sálà mô:nʊ̀.
      1SG with-Kitso(1) S:1PL-FUT-stay:CJ here
      'Kitso and I will stay here.'

    b. *Wena le Lorato lo tlaa apaya dijo.*
      wὲná lí-lʊ̀rátɔ́ ˈlʊ́-tɬáá-àpàjà dì:-dʒɔ́.
      2SG with-Lorato(1) S:2PL-FUT-cook:CJ PL-food(10)
      'Lorato and you will do the cooking.'

As regards gender resolution in constructions with coordinated NPs in subject or object function, if none of the coordinands is a 1st or 2nd person pronouns, it is always possible to apply the following resolution rule (Cole 1955: 429):

– if both coordinands have human referents, *NP₁ le NP₂* governs class 2 agreement: in (15a), both *Leburu* **lì-búrú** and *Lekula* **lì-kúlá** belong to agreement class 5 in the singular and to agreement class 6 in the plural, but *Leburu le Lekula* **lì-búrú lí-lì-kúlá** governs class 2 agreement;

– if both coordinands have non-human referents, *NP₁ le NP₂* governs class 10 agreement: in (15b), *mmidi* **m̀-mídí** 'maize' belongs to gender 3.4, *mabele* **mà-bὲlɛ́** is a plurale tantum of gender 5–6, but *mmidi le mabele* **m̀-mídí lí-má-bὲlɛ́** governs class 10 agreement.

(15) a. *Leburu le Lekula ba tsamaile mmogo.*
      lì-búrú lí-lì-kúlá bá-tsàmáílé m̀mɔ̂:χɔ̀.
      SG-Afrikaner(5) with-SG-Indian(5) S:CL2-leave:PRF:CJ together
      'The Afrikaner and the Indian left together.'

    b. *Mmidi le mabele di jelwe.*
      m̀-mídí ˈlí-má-bὲlɛ́ dí-dʒì:lwé.
      SG-maize(3) with-PL-millet(6) S:CL10-eat:PRF:DJ
      'The maize and the millet have been eaten.'

The semantic basis of this rule is obvious, since the hypernym 'human being' (*motho* **mʊ̀-tʰʊ̀** pl. *batho* **bà-tʰʊ̀**) belongs to gender 1–2, and the hypernym 'thing' (*selo* **sì-lɔ̀** pl. *dilo* **dì-lɔ̀**) belongs to gender 7–10.

---

14 Note that, in Tswana coordinate constructions involving 1st or 2nd person pronouns, the linear order must respect the following hierarchy: 1 > 2 > 3.

According to Cole (1955), an alternative strategy is possible when the coordinands belong to the same agreement class in the plural. In that case, the class in question can be selected instead of class 2 or class 10. According to my own observations, this is possible, but speakers tend to prefer the rule according to which the agreement class assigned to *NP₁ le NP₂* phrases is selected on a purely semantic basis, regardless of the gender of the coordinands. In Tswana, this is the only case in which, within the limits of the clause, semantic agreement takes precedence over morphological agreement.

The resolution rule just formulated raises the following question: what could be the agreement properties of *NP₁ le NP₂* phrases with one of the coordinands human, and the other non-human? In fact, Tswana speakers simply avoid such constructions, and when asked to give a Tswana equivalent of English sentences such as 'The hunter and his dog got lost in the bush', they suggest translations in which the second coordinand in the English sentence is rendered as a comitative adjunct ('The hunter got lost with his dog in the bush').

3.4 Number agreement of verbs

Tswana verbs include an obligatory subject index and optional object indexes. In the first and second person, indexes express person and number, with four possible values: 1st singular, 2nd singular, 1st plural, and 2nd plural. In the third person, indexes express number and gender, with the same possible values as pronouns and adnominals.

(16) Subject indexes in the present affirmative
with first and second person subjects

person	singular	plural
1st	**kì-**	**rì-**
2nd	**ʊ̀-**	**lʊ̀-**

(17) Subject indexes in the present affirmative with third person subjects

gender	singular	plural
1–2	**ʊ́-**	**bá-**
3–4	**ʊ́-**	**í-**
5–6	**lí-**	**á-**
7–10	**sí-**	**dí-**
9–10	**í-**	**dí-**
11–6	**lʊ́-**	**á-**
11–10	**lʊ́-**	**dí-**
14–6	**bʊ́-**	**á-**
17		**χʊ́-**

The paradigm of object indexes expresses exactly the same distinctions.

(18)   First and second person object indexes
person   singular   plural
1st   **ŋ̀-**   **`rí-**
2nd   **χʊ̀-**   **`lʊ́-**

(19)   Third person object indexes
gender   singular   plural
1–2   **mʊ̀-**   **`bá-**
3–4   **`ʊ́-**           **`í-**
5–6   **`lí-**           **`á-**
7–10   **`sí-**          **`dí-**
9–10   **`í-**           **`dí-**
11–6   **`lʊ́-**          **`á-**
11–10   **`lʊ́-**          **`dí-**
14–6   **`bʊ́-**          **`á-**
17   **`χʊ́-**

4 Semantics and discourse

I am aware of no pragmatic function that would be regularly fulfilled by number marking in Tswana. In the same way as in most sub-Saharan languages, deference towards the addressee tends to be expressed at the level of communicative strategies (including physical posture) rather than by the choice of linguistic forms.

The generic use of the second person singular pronoun is possible in Tswana, as well as the use of the third person plural pronoun to encode non-specific reference to a group of humans, but the productivity of these strategies is limited by the remarkably high productivity of impersonal and passive constructions.

As already mentioned, reference to kinds is usually expressed by the plural form of nouns, and the categorization of nouns as count nouns or mass nouns does not show obvious contrasts with European languages. Similarly, plural forms admit non-plural reference in the same way as in the European examples put forward in the questionnaire (for example in questions such as 'Do you have children?').

5 Conclusion

The number category of Tswana shows very few cross-linguistically unusual properties. Its most salient aspects are the morphological complexity of number marking,

the close relationship between number inflection of nouns and gender, the impossibility of dissociating number agreement from gender agreement, and the pervasiveness of gender-number agreement in Tswana morphosyntax. As regards its semantic organization, the number category of Tswana is very similar to the number category of Standard Average European languages.

Abbreviations

APPL	applicative
ATTR	attributive linker
CAUS	causative
CJ	conjoint
CL	agreement class
DEM	demonstrative
DIST	distal
DJ	disjoint
FUT	future
GEN	genitive linker
INF	infinitive
LOC	locative
NP	noun phrase
NUM	numerative linker
O	object index
PL	plural
POT	potential
PRF	perfect
PRS	present
REL	relative
S	subject index
SG	singular

References

Cole, Desmond T. 1955. *An introduction to Tswana grammar*. Cape Town: Longman.

Creissels, Denis. 1997. Prépositions issues de déictiques locatifs: L'exemple du tswana. *Faits de langues* 9. 71–80.

Creissels, Denis. 1999. The role of tone in the conjugation of Setswana. In Jean Blanchon & Denis Creissels (eds.), *Issues in Bantu Tonology*, 109–152. Cologne: Rüdiger Köppe.

Creissels, Denis. 2002. Valence verbale et voix en tswana. *Bulletin de la Société de Linguistique de Paris* 97(1). 371–416.

Creissels, Denis. 2003. Présentation du tswana. *Lalies*. 5–128.

Creissels, Denis. 2004. Non-canonical applicatives and focalization in Tswana. Paper presented at the Syntax of the World's Languages conference. Leipzig, 5–8 August 2004.

Creissels, Denis. 2006. Tswana verb morphology and the Lexical Integrity Principle. *Lingue e linguaggio* 5–1. 49–66.

Creissels, Denis. 2011. Tswana locatives and their status in the inversion construction. *Africana Linguistica* 17. 33–52.

Creissels, Denis. 2013. Les prépositions simples du tswana. In Jesse Teng (ed.), *Prépositions et postpositions: approches typologiques et formelles*, 17–56. Paris: Lavoisier.

Creissels, Denis. 2016. Additive coordination, comitative adjunction, and associative plural in Tswana. *Linguistique et langues africaines* 2. 11–42.

Creissels, Denis. 2017. The conjoint-disjoint distinction in the tonal morphology of Tswana. In Jenneke Van der Wal & Larry Hyman (eds.), *The conjoint / disjoint alternation in Bantu,* 200–238. Berlin/Boston: De Gruyter Mouton.

Creissels, Denis. Forthcoming. Bantu languages: Typology and variation. In Lutz Marten, Ellen Hurst, Nancy Kula & Jochen Zeller (eds.), *The Oxford guide to the Bantu languages*, Oxford: Oxford University Press.

Creissels, Denis, Anderson Chebanne & Heather Nkhwa. 1997. *Tonal morphology of the Setswana verb*. Munich: Lincom Europa.

Doke, C. M. & S. M. Mofokeng. 1957. *Textbook of Southern Sotho grammar*. Cape Town: Longman South Africa.

Marten, Lutz. 2010. The great siSwati locative shift. In A. Breitbarth, C. Lucas, S. Watts & D. Willis (eds.), *Continuity and Change in Grammar*, 249–268. Amsterdam: John Benjamins.

Meeussen, A. E. 1967. Bantu grammatical reconstructions. *Annales du Musée Royal de l'Afrique Centrale* 61. 79–121.

Van der Wal, Jenneke & Larry Hyman (eds.). 2017. *The conjoint / disjoint alternation in Bantu*. Berlin/Boston: De Gruyter Mouton.

II Europe and Middle East

Abdelkader Fassi Fehri

5 Number in Arabic

Abstract: Number as a grammatical category of natural languages is far from being simple. It is not limited to expressing the essential distinction between *singulars* and *plurals*, with their *atoms* and *sums* (or sets) quantity counterparts. As a matter of fact, Arabic contributes a large variety of ingredients and patterns which enrich our understanding of number complexity and diversity. It belongs to the class of *collective-singulative* languages. A derived collective, termed *plurative*, denotes a group individual, or a plurality perspectivized as a *unity*, rather than a multitude. A singulative is formed out of a collective base, and it denotes an individual unity. Atomicities and unities are both needed in the grammar, and they project as atomP and/or unitP, with the right mechanisms. Plurative and singulative derivations characteristically use a second mode of number marking which involves morpho-syntactic convergence of both number and gender. The pronominal number subsystem partially mirrors the nominal subsystem characteristics, but the verbal subsystem is significantly different. To the extent that 'times' can be thought of as verbal classifiers, they only apply to the *n* 'argument'. Other characteristics of the system include dual inflection, human/non-human plurality, double plurality, count-mass phenomenology, etc.

1 Overview

Number as a grammatical category could have been ideally conceived as expressing a distinction of *atoms* and *sums* (or sets) in terms of the quantity they denote, morphologically identified with *singulars* or *plurals*. The picture is obviously more complex in natural languages, since the number system reflects the ways lexical, morphological, and syntactic nominal or verbal categories of the language are organized or classified (as objects, kinds, masses, *Aktionsart* classes, etc.), how it interacts with exact or approximate numerosity, degree (or gradable) quantity, or mereological relations among sets, or how various singularities or pluralities are conceived or perceived. Arabic contributes a large variety of ingredients and patterns that can enrich our understanding of number, and the various ways it is expressed in the various languages. It belongs to the class of *collective-singulative* languages, where a singular is formed out of a 'collective' by singulative morphology, in addition to the class of the more common *singular-plural* languages like English. The first basic patterns of number representing the two inflectional number 'types' are illustrated in (1) and (2), respectively:

(1) a. *muslim-un* → *muslim-uu-na*
       muslim-NOM   muslim-PL-NOM
       'A Moslem'   'Moslems'

b. *kalb-un*   → *kilaab-un*
   dog-NOM   dog.PL-NOM
   'A dog'   'Dogs'

(2) *baṭṭ-un*   → *baṭṭ-at-un*
   duck- NOM   duck-UNIT-NOM'
   'Ducks; one or more ducks'   'A/one duck'

In (1a), the inflectional pluralization is obtained by the addition of a long vowel suffix expressing plurality, having the singular form as the source of derivation. In (2), by contrast, it is the singular (or rather the so-called 'singulative') which is derived from a nominal base expressing a general number, and denoting either singular or plural (although morphologically unmarked as such). The alternations in (1a) and (1b) add a morphological complexity, namely a distinction between concatenative plurals (or so-called *sound* plurals in the Arabic tradition) in (1a) and non-concatenative plurals (or so-called *broken* plurals) in (1b), depending on whether the plural is suffixal or infixal. We can see then that when English appears to treat *dog*, *duck*, and *moslem* equally as singulars, and *dogs*, *ducks*, and *moslems* equally as plurals, Arabic makes more subtle distinctions across the board, comparable in part to those found in languages as diverse as Persian, Chinese, Burmese, Welsh, etc.[1] It is important to note that the singular/singulative in (2) is formed via a homophonous morpheme to the feminine, which is best treated as forming a *unity* (or a unitizer), being a sort of classifier, rather than a traditional gender.

There are also collectives in the language that denote a group of entities, which can be best analysed as *unities*, constructing in fact a unity over a plurality in the grammar. I call them pluratives, because unlike plurals in (1), they use a suffix homophonous with the feminine to denote groups. The basic pattern is exemplified in (3):[2]

(3) *najjaar-un*   → *najjaar-at-un*
   carpenter-NOM   carpenter-UNIT-NOM
   'A carpenter'   'Carpenters (as group)'

---

1 Wiese (2012) observes that in classifier languages like Chinese or Japanese, object denoting nouns are collectives as a rule. It is also the case in Persian, according to her analysis, where the singular-singulative is morphologically marked, having a collective as a base:

(i) *ketaab* 'book'; "books; one or more books" → *ketaab-**ii*** 'book– unit'; "a/one book"

For more variation across languages, see Fassi Fehri (2018, Ch. 5).

2 Note that *najjaar-at-un* can also (ambiguously) mean the feminine of *najjaar* (the female carpenter). But this interpretation is irrelevant to our discussion. See Fassi Fehri (2018, Ch. 2) for the multiple roles of Gender.

One essential feature of the system then is the morpho-syntactic convergence of number and gender in a number of patterns, and their significant interaction. The pronominal system partially mirrors the nominal system, but the verbal system is significantly different. Verbs cannot be headed (or selected) by a classifying gender (or individual phrase, 'DivP'), Number (or ≠P), or *n* (nP). 'Verbal classifiers' never apply strictly to verbs, but only to the event argument or event phrase. To the extent that 'times' can be analysed as verbal classifiers, they apply only to the *n* 'argument'. If so, then Number is not 'verbal' in any proper sense.

2 Pronominal, nominal, and verbal number

2.1 Generalities

The present contribution is especially dedicated to investigating two *modes* of number marking in Arabic: the common singular-plural and the less common collective-singulative. The two modes co-occur and interact in shaping the profile and patterns of the Arabic number system, be they pronominal, nominal, or verbal. The term 'Arabic' is meant to cover the various Arabic varieties, including Modern Standard Arabic (which is chiefly described here), and the colloquials (when they are specifically referred to, e.g. Moroccan Arabic). The facts described here are hardly controversial or questionable throughout authoritative references. Moreover, these patterns are instantiated in quite all the varieties spoken throughout the Arab world (the west and the east Maghreb and Machreq 22 countries), as well as the Moslem world, or other countries with non-natives or heritage speakers (recent statistics from *Ethnologue* point to 274 million Standard Arabic speakers, and those from UNESCO to more than 422 million for the different varieties). What is original in both the description and the analysis is how the various patterns are integrated into one interacting system.

Number marking is obligatory in general, and it normally co-occurs with the marking of gender of the singular, as illustrated clearly by the dual form in (4):

(4) *qiṭṭ-un*   → *qiṭṭ-**at**-un*   → *qiṭṭ-**at**-**aa**-n*
     cat-NOM   cat-FEM-NOM   cat-FEM-DUAL-NOM
     'A male cat'   'A female cat'   'two female cats'

Here the feminine *qiṭṭ-**at*** is derived from the masculine *qiṭṭ* by suffixing the feminine marker −*at*, and the dual of the feminine adds the suffix −*aa*, illustrating the fact that gender and number co-occur autonomously on the unmarked base.

In the plural case, however, two apparent important exceptions to this rule need to be dealt with: (a) the group plural that I call the plurative, and (b) the non-human plural. These two forms of pluralities are characterized by a sort of syncre-

tism or convergence of number and gender, according to which the vowel lengthening characteristic of pluralization is absent, and gender (to express sex) is no longer variable. The feminine used in these two cases is instead interpreted as 'unitizing' or grouping the plurality. In the construction (5), the plurative noun phrase is morphologically marked with the unit affix *–at*, and the verbal predicate agrees with the subject by using the same unit marker (rather than the more common plural *–uu*):

(5) *l-qatal-**at**-u*   *xtabaʔ-**at***
     The-killer-UNIT-NOM hid-UNIT
     'The killers (as group) hid.'

In such a construction, the only interpretation available is that of a collective group and action (rather than a distributive plural or action).[3]

The second 'exception' to the co-occurrence of both gender and number on the noun and/or the predicate is that of non-human pluralization. In (6), the broken plural does not show any mark of gender (of the singular), and the predicate cannot agree by using *–uu* (the normal mark of the plural), the only mark available being the uniform feminine ('plural'):

(6) *l-kilaab-u*   *nabaḥ-at*   *(* nabaḥ-uu)*
     the-dogs-NOM barked-FEM (barked-PL)
     'The dogs barked.'

I will return to constructions involving this kind of syncretism below.[4]

---

3 Fassi Fehri (2018, Ch. 2) has detailed the non-canonical roles played by the feminine *–at* in Arabic, including the classifier role expressing unity.
4 Rijkhoff (2002) observes that Burmese has sortal singular and collective classifiers which occur obligatorily with numerals, to count two kinds of discrete entities, singulatives and collectives. The following example illustrates the collective case:

(i) pyà hnǎ ouñ
     bee two swarm
     'Two swarms of bees'

Likewise, Fassi Fehri and Vinet (2004) note that there is a collective group interpretation characteristic of the Chinese suffix *-men*, rather than a normal plural one, as in (ii):

(ii) Xiǎo Qiángmen
     Xiao Qiang-MEN
     'Xiao Qiang's group'

2.2 Pronominal number

In this subsection, I concentrate on the number of personal pronouns, leaving aside so-called demonstrative, relative, indefinite, or interrogative pronouns (the number values of which mirror only partly those of personal pronouns). Pronouns normally vary in person (1, 2, 3/Ø), number (plural, dual, singular), gender (feminine, masculine), and case (nominative, accusative, genitive), but syncretisms or gaps in paradigms are also found. Values of pronominal number also include *plurative* and *singulative*, seen as unities, in contrast with normal singulars or plurals, which express overtly no unity, being normally interpreted as distributive. There is also a human/non-human distinction which is manifest only in the expression of pluralities. The essential distinction in terms of form is between (a) 'free' or 'independent' pronouns (so-called *munfaṣil* in the Arabic tradition), which exhibit both morphological and syntactic autonomy, and (b) 'bound' or 'clitic' pronouns (so-called *muttaṣil*), which have no such autonomy, although both forms are marked typically for person. The Arabic tradition also identifies a third class, called 'hidden' (*mustatir*), which is typically characterized by the absence of manifestation of person, although there may be marking of number and gender. This case can be clearly analysed as a case of so-called Pro Drop, although bound pronouns in (b) may or may not be so analysed, being instead cases of pronoun incorporation. Person and case aside, number (and gender) values of pronouns reflect those of nouns, although some complexities and subtle distinctions found in nouns are not found in pronouns, including 'double plural', 'plural of abundance', 'paucity', etc. Moreover, no verbal plurality feature is reflected in pronominal number.[5]

Free pronouns normally occur in nominative positions where they have no immediate governor they are forced to incorporate to, as when they are subjects of copulatives, dislocated topics, etc. Table 1 provides the list of these pronouns. Their internal structure consists of (a) a morphological support like *ʔan* (specific to first and second person, and vanishing with bound forms), (b) a person indicator in the first suffix position (*t* or *h*), and (c) a number indicator in the second suffix position (-*aa*, or −*uu*; sequential or split ordering being more transparent in 2nd and 3rd person than in 1st).

The following examples illustrate some forms and contexts of nominative free pronouns:

(7) *ʔanta   masʔuul-un*
    you.NOM responsible-NOM
    'You are responsible.'

---

5 For detail on properties of Pro drop and pronoun incorporation, see Fassi Fehri (1993, 1996–2000), and the references therein.

Tab. 1: Free nominative pronouns.

Pers		Num		
		Sg	**Dl**	**Pl**
1		*Ɂan-aa*	*na-ḥnu*	
2	masc	*Ɂan-ta*	*Ɂan-tum-aa*	*Ɂan-tum (-uu)*
	fem	*Ɂan-ti*		*Ɂan-tun-na*
3	masc	*huwa*	*hum-aa*	*hum (-uu)*
	fem	*hiya*		*hun-na*

Tab. 2: Bound nominative pronouns.

Pers		Num		
		Sg	**Dl**	**Pl**
1		*katab-**tu** **Ɂ**-aktub*	*katab-**naa***	***n**-aktub*
2	masc	*katab-**ta** **t**-aktub*	*katab-**tum-aa*** *t-aktub-**aa***	*katab-**tum**(-uu)* *t-aktub-**uu***
	fem	*katab-**ti** **t**-aktub **ii***		*katab-**tun-na*** *t-aktub-**na***
3	masc	*kataba* ***y**-aktub*	*katab-**aa*** ***y**-aktub-**aa***	*katab-**uu*** ***y**-aktub-**uu***
	fem	*kataba-**t*** ***t**-aktub*	*kataba-**t-aa*** ***t**-aktub-**aa***	*katab-**na*** ***y**-aktub-**na***

(8) *naḥnu   ḥadar-naa*
　　 we.NOM present.PAST-1.PL
　　 'We were present.'

Table 2 provides the list of bound nominative pronouns (in bold), specifically when the pronominal subject is incorporated into the inflected verb, conjugated in the perfect/past or the imperfect/present, respectively. The form of the pronoun is split or discontinuous in the imperfect, but continuous in the perfect, yet the two positions of person and number are essentially preserved in both conjugations. It is striking that *h-* disappears in third person hidden pronouns, but number (and gender) are preserved.[6]

---

6 The relevance of the two positions has been argued for in Fassi Fehri (1996–2000).

Tab. 3: Free non-nominative pronouns.

Pers		Num		
		Sg	Dl	Pl
1		*ʔiyyaa-ya*	*ʔiyyaa-naa*	
2	masc	*ʔiyyaa-ka*	*ʔiyyaa-kum-aa*	*ʔiyyaa-kum*
	fem	*ʔiyyaa-ki*		*ʔiyyaa-kun-na*
3	masc	*ʔiyyaa-hu*	*ʔiyyaa-hum-aa*	*ʔiyyaa-hum*
	fem	*ʔiyyaa-haa*		*ʔiyyaa-hun-na*

Tab. 4: Bound non-nominative pronouns.

Pers		Num		
		Sg	Dl	Pl
1		*-ya* (*-i*)	*-naa*	
2	masc	*-ka*	*-kum-aa*	*-kum*
	fem	*-ki*		*-kun-na*
3	masc	*-hu* (*-hi*)	*-hum-aa* (*-him-aa*)	*-hum* (*-him*)
	fem	*-haa*		*-hun-na* (*hin-na*)

Table 3 provides a list of free non-nominative pronouns. Here, too, we can see that the essential of the internal pronominal structure is preserved, modulo the difference in the morphological support, and the difference in case.

Here is an illustration:

(9) *ʔiyyaa-ka   qaṣad-tu*
    you.ACC   addressed-I
    'It is you that I addressed.'

Table 4 provides a list of bound non-nominative pronouns. The list is identical to that found in 3, modulo the absence of the support *ʔiyyaa*.

As far as number is concerned, observe that Pers 1 pronouns have only two number forms, a singular and a plural (or non-singular). No distinction is made between dual and plural, and there is no distinction in terms of gender either. Pers 2 and Pers 3, in contrast, distinguish three numbers (singular, dual, and plural). The dual

is neutral with respect to the gender distinction. In Pers 2, the dual forms *-tum-aa* or *-kum-aa* differ in terms of case. In the third person form, only *-aa* is present, and Pers 3 is hidden. Similar observations can be made with respect to the plural, except that a gender distinction is made: *tum/tunna, hum/hunna*. In Pers 3, the person is also hidden, which results only in a number/gender realization: basically *-uu* in the masculine, and *-na* in the feminine. In the independent forms, the singular Pers 3 *huwa* and *hiya* are distinguished by gender. Gendered singular plays an important role in marking specific pluralities, specifically non-human plurals, and pluratives. Then any form of Pers 3 feminine singular can be used in the appropriate context to track these pluralities. In construction (6) above, the *-at* is used to backtrack its non-human plural antecedent, while the single feminine pronoun is foregrounding the human plural (perspectivized as a group) in (10):

(10) *hiya r-rijaal-u*
      she   the-men-NOM
      'That is men.'

In (10), *hiya* is used in place of the plural *hum* as the appropriate anaphor for *rijaal*, seen as a group. Observe that the pronominal number subsystem is derivative from the nominal number subsystem in terms of morphology, since it represents one sub-set of it. But there is no distinctive morphology in pronominal forms that provides the traditional sound/broken plurality in nouns, nor a morphology of plurative pro-noun formation as distinct from plurative noun formation.

2.3 Nominal number

Arabic nouns (as they surface in clauses) are traditionally thought to be 'numbered', in the sense that they are obligatorily associated with one of the three values of number. The dual and the plural are morphologically marked, while the singular is unmarked. Contrastive values are illustrated in (11):

(11) a. *fii d-daar-i   rajul-un*
         in the-house-GEN man-NOM
         'There is a man in the house.'

      b. *fii d-daar-i   rajul-aani*
         in the-house-GEN man-DUAL.NOM
         'There are two men in the house.'

      c. *fii d-daar-i   rijaal-un*
         in the-house-GEN men-NOM
         'There are men in the house.'

The relevant interpretation in (a) is that there is 'one' man (and no more), in (b) there are 'two' men (and no more), and in (c) there is more than 'one', or more than 'two' (by contrasting (c) with the dual in (b)). These nouns denote either single atoms or sums (or sets) of atoms, but they are not set free in denoting atoms or sums. In addition to these nominal forms specified for values of number, we have to add *general* nouns like *samak* 'fish' (sometimes called 'collectives'), which can be considered as vaguely numbered or atomistic, denoting one or more objects. General nouns can be reasonably distinguished from substance *mass* nouns like *zayt* 'oil', which are neither numbered nor atomic in any obvious sense. The latter differ from the former in that they cannot be identified with atoms or sums of atoms. Note that singulars or plurals are not always associated with fixed quantity semantics or interpretation (that is, 1, > 1, or > 2), in contrast e.g. to duals.[7]

There are various forms of singularities or pluralities (in addition to their rich allomorphy), depending on the various classifications of nouns, their syntax, and their interpretation. Thus, some plurals are quantity phrases formed in syntax through the projection of NumP (or #P in Borer's 2005 terms), while other pluralities are 'lexical' or *n* plurals in the sense of Acquaviva (2008). Thus, although *rakb* 'riders' or *ṣaḥb* 'companions' are 'born' as plural, in the sense that they denote only sums (of atoms), there is presumably no standard way to derive them from their singular counterparts (*raakib* 'rider' or *ṣaaḥib* 'companion'), and they should be treated as lexical or *n* plurals. Likewise, it is possible to treat nouns like *kalb* 'dog' as a 'lexical' singular, denoting a singular atom, under *n*, rather than a quantificational # singular. But at least a singular noun like *samak-at* 'fish-unit' is formed through singulative morphology, more plausibly in syntax, via a unitizer phrase (UnitP, as in Fassi Fehri 2018, Ch. 5, or equivalently a divider phrase, DivP, as in Borer *ibid*). If true, then (semantic) number is not confined only to NumP; it can be generated lower in the structure (nP, UnitP, or other potential positions associated with root phrases, or RootP), as well as higher in determiner phrases (or DP). See Section 3 below for the detailed syntactic architecture of traditional noun phrases.

Nominal plural morphology (which often extends to adjectives) varies with respect to whether it is a concatenative 'sound' suffix (added to the stem), as 'external plural', or it is a non-concatenative 'broken' affix (involving internal vowel change infixation), also called 'internal plural'. The two forms of number are exemplified through the sound/broken plural doublet given in (12):

(12) a. *mudiir-un* → *mudiir-uu-na*
     director-NOM   director-PL-NOM
     'A director'   'Directors'

---

7 On general nouns, see Corbett (2000: 9–19). These are equivalent to *set* nouns (Rijkhoff 2002), or *kind* nouns (Fassi Fehri 2004). See also subsection 4.1 for semantic qualifications.

b. *mudiir-un* → *mudaraaʔ-u*
    director-NOM   director-PL.NOM
    'A director'   'Directors'

The dual morphology, by contrast, is always concatenative/sound, as exemplified by (4) and (11b) above.

There is a third form of the plural, often neglected, although quite productive, which I call the plurative. It is characterized by its feminine suffix ending, which makes it essentially a special form of sound plural, although it often mixes the two morphologies (in largely unpredictable ways). Examples of pluratives (denoting groups) are given in (13a) and (13b):

(13) a. *ḥaddaad*   → *ḥaddaad-**at***
        'blacksmith'   'blacksmiths (as group)'
        majuusiyy   → *majuusiyy-**at***
        'magian'   'magians (as group)'
    b. *baaʔiʕ*   → *baaʕ-**at***
        'seller'   'sellers (as group)'
        saadin   → *sadan-**at***
        'guardian'   'guardians (as group)'
        ṭaaġ-in   → *ṭuġ-**aat***
        'tyrant'   'tyrants (as group)'

Triplets or doublets of plurals are often correlated with some subtle or specific meaning differences, but they are also interchangeable forms in many other cases. When distinctions between various unmarked or marked singulars enter the picture in forming plurals of various kinds, the syntax of number becomes necessary to distinguish lexical or *n* plurals from *number* or # plurals, *head* and *modifier* plurals, or plurals of *atomicities* and plurals of *unities,* etc. (see Fassi Fehri 2018, Ch. 5 for detail).

In practice, the sound plural appears to have a rather limited distribution, compared to the broken plural (or the plurative), and they are often in complementary distribution, rather than competition. As McCarthy and Prince (1990: 212) observe: "… essentially all canonically-shaped lexical nouns of Arabic take broken plurals", while the sound plural is "systematically found only with … proper names, transparently derived nouns or adjectives such as participles, deverbals and diminutives; non-canonical or unassimilated loans". Contrary to the (sound) dual number, which basically applies to any noun or adjective, the sound plural is restricted to a set of nominal forms that must meet formal and syntactic criteria (including the limit to masculine nouns referring to rational beings). And although it is "qualitatively productive", as Boudelaa and Gaskell (2002: 326) put it, broken pluralisation is also qualitatively productive.

Syntactically and semantically, all the three forms of plurals can provide bases for generating number under #P, the locus of quantity phrases. To illustrate:

(14) *ḥadara ṭ-ṭullaab-u*
came   the-students-NOM
'The students came.'

(15) *fii l-madrasat-i mudarris-uuna   ʔakfaaʔ-u*
in school-GEN   teachers-NOM-ACC competent.PL-NOM
'There are competent teachers in the school.'

(16) *ttafaqa s-sadan-at-u*
agreed the-guardians-UNIT-NOM
'The guardians agreed.'

Moreover, all three forms appear in contexts where they are presumably generated lower in the structure (under nP or classifier phrase, ClP). For example, the three forms appear equally as a complement of a low numeral, as in the following constructions:

(17) *ṯalaaṯ-at-u   ṭullaab-in*
three-FEM-NOM students-GEN
'Three students'

(18) *ṯalaaṯ-u   ṭaalib-aat-in   yamaniyy-aat-in*
three-NOM student-FEM.PL-GEN yemeni-FEM.PL-GEN
'Three female Yemeni students'

(19) *ṯalaaṯ-at-u   sadan-at-in*
three-FEM-NOM guardian-GROUP-GEN
'Three guardians'

The form of the plural on the nominal is required for agreement purposes with the numeral (Krifka 2013). Other uses of plural forms are good candidates to be generated under *n* (see e.g. Acquaviva 2008, Fassi Fehri 2004, 2012 Ch 5, and references therein).

Attempts have been made to differentiate the concatenative/non-concatenative (sound/broken) plural forms by associating them with distinct syntax and meanings. For example, one apparent descriptive characteristic of the broken plural is its lack of (overt) morphological gender (of the singular), compared to the external sound plural noun (which normally manifests gender), as the contrast between (20) and (21) illustrates:

(20) *kaatib-at* → *kuttab*
     writer-FEM   'writers (undifferentiated in gender)'

(21) *kaatib-at* → *kaatib-aat*
     writer-FEM   writer-FEM.PL; 'female writers'

This line of thought, however, is contradicted by the behaviour of broken plurals with (low) numerals. With the latter, the numeral varies in gender (in a 'polar' manner) depending on the gender of the counted noun. This is clearly seen in the variation in gender already found in (17) to (19) above, where the numeral agrees (in a reverse way) with the singular (rather than the plural) in gender. If so, then the broken plural is not devoid of gender. Moreover, the broken plural cannot be assimilated only to a 'gender' or 'classifier', leaving only the sound plural to be Num or #. The situation gets more complex when the plurative enters the picture. A finer syntax of broken plurals can be shown to be *flexible*, in the sense that they can be generated under *n* or *Num*. See Section 3 below for more detail.

The singulative functions most often as an 'individuator' when it applies to a general noun, or as a 'packager' when applied to mass nouns (see Fassi Fehri 2018, Ch 5 for detail). Both singulatives require an extra projection in syntax to be formed (where a classifier is introduced). As for the plurative, it can have two semantic facets. On the one hand, it is just another form of morphological plural inside the set of sound and broken forms, yielding a sum plural. But its most specific narrow semantic sense becomes clear when it is invariably associated with feminine singular in morpho-syntax. It then triggers unit agreement, and it occurs only with collective events. It significantly contrasts with both regular sound or broken plurals that are typically plural in morpho-syntax, vary in gender, and are interpretable as sums. The plurative has no true sum interpretation; rather, it is a *unity* of the various atomicities included in the sum, a whole which has its own characteristics, and it has no variable gender (see Fassi Fehri *ibid*, and Meirav 2003 for more on 'unities').

Plural of mass is also productive. The following pair of examples illustrate two of its uses:

(22) *fii n-nahr-i   t-ajrii   miyyaah-un kaṯiir-at-un*
     in the-river-GEN FEM-flow waters-NOM many-FEM-NOM
     'In the river, a lot of water flows (many waters flow).'

(23) *ḏuq-tu ṯalaaṯ-at-a   zuyuut-in*
     tasted-I three-FEM-ACC oils-GEN
     'I tasted three (kinds of) oils.'

In (22), the plural *miyyaah* is interpreted as an 'abundant plural', meaning a big quantity of water, rather than as an atomic plural. It is not a classifier (being neither

an individuator nor a packager). It is not a Number phrase either, because it does not yield a sum interpretation. It is more appropriately treated in the grammar as a modifier adjunct. In (23), the plural is a taxonomic plural (denoting the number of kinds, rather than individuals).

Double plurals can be morphologically formed from an already morphological plural, as is illustrated in (24):

(24) *fii l-intixaabaat-i   xuruuq-aat-un   kaṯiir-at-un*
     in the-elections-GEN violations-PL-NOM many-FEM-NOM
     'In the elections, there are many violations.'

In this construction, *xuruuq* is already a (broken) plural of *xarq*, and *xuruuq-aat* is a concatenative double plural (based on the former). But when such a morphologically double pluralization occurs, its interpretation is not 'a plural of plural' *per se*, but rather a 'plural of abundance', meaning 'a lot', and the second plural acts in fact as a modifier of the first, focusing on the important quantity (of violations). See Section 3 below for detail.

Within the category of 'plural', Arab grammarians also traditionally draw a distinction between a plural of paucity (*jamʕ qill-at*, a 'plural of few', claimed to be referring to a quantity from three to ten), and a multal plural (*jamʕ kaṯr-at*, a 'plural of multitude'), referring to a big quantity, or more than ten (Wright 1971, Ojeda 1992). The pair in (25) illustrates the intended contrast:

(25) a. *kalb* → *ʔaklub*
      'dog' 'a few dogs'

    b. *kalb* → *kilaab*
      'dog' 'many dogs'

But such a distinction (which applies only to broken plurals) is rather illusory in practice as far as Modern Standard Arabic is concerned, and to what extent it has ever been real in the past is still undecided (see Ferrando 2006).

2.4 Verbal number

Verbal number can be confused with neither verbal agreement, nor verbal aspect. To characterize verbal number is part of the general understanding of what number is, although specific numbering properties of verbs or events can be identified. From early literature (Boas 1911, Cusic 1981, Durie 1986, Mithun 1988), it was understood that verbal number is 'special'. It has no clear parallel to the meaning 'more than one', no dual or trial, no specific quantity or measure or cardinality that can be associated with the bare verb or vP. Verbal plurals are then only 'special plurals'

(Souckova 2011, Veselinova 2013, Doetjes 2008). Verbal number could even be only 'derivational', as Corbett (2000: 243–9) put it. It is true that the event can be counted by using event units, or counting the number of times, but these constituents are best seen as modifiers of the verbal event, rather than part of its extended projection, projecting some NumP as in the case of nominals. However, the parallelism is still worth exploring for more general concerns.[8] For example, verbs (or events) can be born as plurals 'in the lexicon' (Kratzer 2008). An activity verb like *raqaṣa* 'dance' can denote a plurality of dancing events (in addition to a singular event). More precisely, the event noun *raqṣ* can be conceived as a counterpart of a general noun like *samak*, naming a *kind event*. It denotes one or more events from the start, in the form of the cognate event in (26). Alternatively, one can measure the number of cognate *event units*; they can be 'one', or 'more than one', as in (27):

(26) *raqaṣa r-rajul-u   raqṣ-an   ġariib-an*
     danced the-man-NOM dancing-ACC strange-ACC
     'The man danced a strange dancing.'

(27) *raqaṣa r-rajul-u   ṯalaaṯ-a   raqaṣ-aat-in*
     danced the-man-NOM three-ACC dance-UNIT.PL-GEN
     'The man danced three dances.'

Also, one can measure the number of *times* the event occurred, as in (28):

(28) *raqaṣa r-rajul-u   ṯalaaṯ-a   marr-aat-in*
     danced the-man-NOM three-ACC times.PL.FEM-GEN
     'The man danced three times.'

It is possible to combine measurements of occurrences of the event and measurements of times, as in (29):

(29) *raqaṣa r-rajul-u   ṯalaaṯ-a   raqaṣ-aat-in   ṯalaaṯ-a*
     danced the-man-NOM three-ACC dance-UNIT.PL-GEN three-ACC
     marr-aat-in
     time-PL.FEM-GEN
     'The man danced three dances three times.'

---

8 I concur with Souckova (*ibid*) in distancing verbal plurality or pluractionality from both Aktionsart (contra e.g. Cusic 1981, Wood 2007) and situation aspect (contra e.g. Van Geenhoven 2004). Reduplication, intensive, event plurality, or participant plurality are then a matter of number, rather than temporal interpretation. See Fassi Fehri (2012, Ch 11), and Cabredo Hofherr and Laca (2012) for more detail.

That results in nine dances altogether. Moreover, distributive, collective, iterative, or cumulative interpretations can arise with these verbs, which fall under Krifka's (1992: 33–36) *cumulative universal*, according to which simple predicates in natural language are typically cumulative.

2.4.1 Pluractional morphology

In Arabic, productive pluractional morphology involves consonant gemination, reduplication, or vowel lengthening.[9] These internal morphological operations apply to basic roots, to form complex roots, inducing various pluractional interpretations, including so-called repetitive/intensive, 'attenuative' readings, or 'interaction/participation' (see Greenberg 1991, Fassi Fehri 2001, 2003 for varieties of these meanings). Consider (30), in which the medial glide of the verb is geminated, compared to (31), the simple verbal form:

(30) *jawwala   r-rajul-u   fii l-bustaan-i*
     walked.INTENS the-man-NOM in the-garden-GEN
     'The man took a lot of walks in the garden.'

(31) *jaala   r-rajul-u   fii l-bustaan-i*
     walked the-man-NOM in the-garden-GEN
     'The man walked in the garden.'

In (30), the event of walking is said to be repetitive or intensive, denoting many events of walking (or a lot of walks). But the simple form in (31) is compatible with one event of walking taking place. The difference in interpretation is clearly captured by the following contrast:

(32) a. *jaala   r-rajul-u   jawl-at-an   waaḥidat-an*
        walked the-man-NOM walk-UNIT-ACC one-ACC
        'The man walked one walk.'

   b. **jawwala   r-rajul-u   jawl-at-an   waaḥidat-an*
       walked.INTENS the-man-NOM walk-UNIT-ACC one-ACC

The ungrammaticality of (32b) is due to the incompatibility of the plurality of the event with the single cognate event.

---

9 The term *pluractionality*, originally introduced by Newman (1990, 2000), is intended to highlight the specificity of verbal number. But *event number* appears to be a more convenient and more general term, which facilitates the comparison of the various specific numbers.

A second form of event plurality is obtained through full reduplication of biliteral roots to form quadriliteral forms, with a subevent interpretation, as in the following contrast:

(33) *xanna* → *xanxana*
  'to speak through the nose, nasalize' 'to nasalize smoothly and repeatedly'

In its simple form, the biliteral base of the root denotes the whole event, while in its complex form, it is reiterated and denotes only subevents of the whole, leading to an 'attenuative' interpretation. It is hard to imagine that this form of verb or others (such as *qahqaha* 'to giggle', *ṭaqṭaqa* 'to patter', *zaqzaqa* 'to chirp', or any other onomatopoeic forms) obtain only through singular action. Rather, plurality of events (or subevents) is obligatory.[10]

2.4.2 Event and participant readings

Intensive transitive constructions can give rise to double readings, where the plurality is understood as either involving a plural event, or a plural participant. Thus in construction (34), where both participants are singular, only the event reading is available. But in (35), both readings are found:

(34) *jarraḥa*   *l-jundiyy-u*   *ṭ-ṭifl-a*
   wounded.INTENS the-soldier-NOM the-child-ACC
   'The soldier inflicted many wounds to the child.'

(35) *jarraḥa*   *l-jundiyy-u*   *l-ʔaṭfaal-a*
   wounded.INTENS the-soldier-NOM the-children-ACC
   (a) 'The soldier inflicted many wounds to the children.'
   (b) 'The soldier wounded many many children.'

Keeping in mind the essentials of Fassi Fehri's (2012, Ch 11) analysis, though with some refinements, I propose the following structures for the two readings (*e* for event; and the broken PL is merged with the nP):

---

10 See Glanville (2018) for more detail on these forms.

(36) *Event reading*

```
                eP
              /    \
            e        DP
          /  \      /   \
       [PL]   e    D     NumP
         |    |          /    \
     jarraḥa        Num      nP
                      \      /
                    l-ʔaṭfaal-a
```

(37) *Participant reading*

```
                eP
              /    \
            e        DP
          /  \      /   \
       [PL]   e    D     NumP
         |    |         /     \
     jarraḥa        NumP      NumP
                      |       /    \
                    [PL]   Num     nP
                             \     /
                          l-ʔaṭfaal-a
```
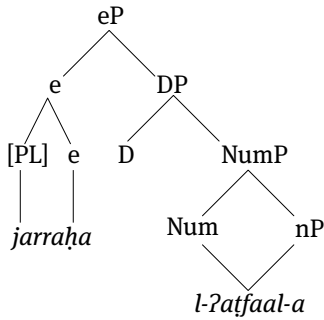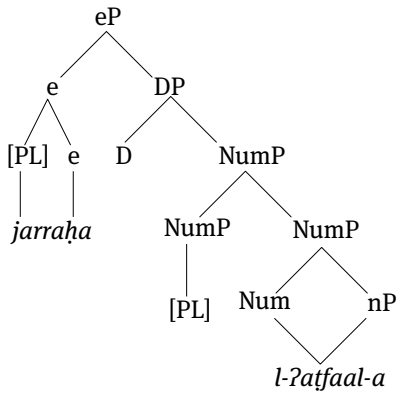
In the event reading, the morphological plural is interpreted on the verbal root head, represented by *e* (for event, instead of *v*), and the intensive event reading results from adjoining the PL to *e*, as (an adverb-like) modifier. Modifying a (lexically born) plural root by a PL gives the effect of plural intensification, by meaning composition. In the participant reading, the 'double pluralization' applies to the DP, increasing the quantity of objects, and (only indirectly) that of events. What is novel is that the mechanism which accounts for differences in interpretations is semantic composition, rather than Agree, as clarified in Section 4. This choice should come as no surprise since verbal plurality has been kept separate from verbal agreement from the start.[11]

---

11 See Section 4 for more precise discussions. In fact, the interpretable Pl is adjoined to *e* in (36), and to Num in (37). Being a modifier in both cases, there is a theta identification that applies to the two plurals. I assume that the Pl on *e* in (37) is not interpretable. See Chomsky (1995) on (non-)interpretability of features, and Wiltschko (2008) on the head/modifier status of Pl, as well as the logic of reasoning.

Note that both event plural and participant plural readings are involved in interpreting so-called 'participative verbs' (*ʔafʕaal l-mušaarak-at*), as in the following construction:

(38) *raaqaṣa   r-rajul-u   l-marʔat-a*
    danced.PL the-man-NOM the-woman-ACC
    'The man took part in dancing with the woman.'

Here pluractionality is marked via vowel lengthening. The interactional verb requires a plural participant, but the two participants are distributed in different positions. The event is also plural because each dancer is supposed to dance with the other. Likewise, reciprocal verbs like those in *ta-qaatala* 'to kill each other', *ta-šaatama* 'to gossip about each other', are also participative verbs, and hence involve both plurality of events and plurality of participants.[12]

3 Agreement and the syntax of number

Agreement in number with verbal and adjectival predicates is obligatory and regular. It exhibits a rather limited number of forms of exponence, compared to the richness of allomorphy in nominal number. Thus, there is no equivalent in the verbal domain to the nominal broken plural allomorphy (which counts more than thirty-five forms, including paucal or multal), although there are a few that occur on adjective predicates. The regular pattern of predicate agreement is then sound, or concatenative. For example, a human plural triggers a regular plural agreement on the verb. For the masculine plural, it takes the form of a long vowel *−uu*, and for the feminine plural, it is *−na*, as in the following pair of constructions:

(39) *l-ʔawlaad-u   katab-uu d-dars-a*
    the-boys-NOM wrote-PL   the-course-ACC
    'The boys wrote the course.'

(40) *l-banaat-u   katab-na   d-dars-a*
    the-girls-NOM wrote-PL.FEM   the-course-ACC
    'The girls wrote the course.'

This form of plural is normal or regular in the sense that it can apply in most contexts. It is also regular in that the plural affix manifests the gender of its singular. I assume that the agreement operation is a form of *Agree* in Chomsky's (1995) sense.

---

12 See Fassi Fehri (2001; 2012, Ch 11) and references therein for detail.

Two forms of plural agreement differ from the regular one described here: (a) the non-human plural agreement (as in (6) above), and (b) the plurative (as in (10) above). Here, there is no overt plural morphology, and no marking of the gender of the singular. Rather, the plural form is invariably feminine singular. I will attempt an explanation of differences, and how they are reflected in the Agree operation.

3.1 Agreement patterns: pronominal and nominal

Consider first the three forms of the broken plural subject in (41). They have only one verbal form counterpart, carrying a feminine (singular) suffix as a mark of non-human plural:

(41) *l-ʔusd-u*   / *l-ʔusuud-u*   / *l-ʔaasaad-u*   *zaʕar-at*
      the-lions-NOM the-lions.many-NOM the-lions.many.many-NOM roared-FEM
      'The (many) lions roared.'

It is worth noting that the distinct plural allomorphy of the nominal subject is not paralleled by a distinct agreement allomorphy on the predicate (which would have taken the controller formal differences into account). Moreover, the agreement in phi features is not regular in the sense that it does not manifest overt plural number and overt gender.

Consider now a third pattern of plural agreement which I call *plurative* agreement.[13] This pattern is not driven by any human/non-human distinction. Its core case is when a plural nominal is formed as a plurative via the unit suffix -*at*, to denote a plural *unity*; it then triggers copying the unit feature on the verbal predicate. Thus *qatal-at* 'killers as group' (plural of *qaatil* 'killer') triggers unit agreement on the verb predicate, as in (5) above, repeated here as (42):

(42) *l-qatal-at-u*   *xtabaʔ-at*
      the-killer-UNIT-NOM hid- PL
      'The killers (as group) hid.'

In a sense, 'copying' the feminine unity feature on the predicate results in a unity (or group) interpretation of the event as well.

But note that the plurative DP can also occur in contexts where the verbal predicate carries a regular plural. In the context (44), the event is interpreted as distribu-

---

13 I avoid calling it a 'group' agreement because groups do not behave alike with respect to this form of agreement. See Fasssi Fehri (1988, 2004), Ojeda (1992), Zabbal (2002), Acquaviva (2008), among others, for discussion of Arabic group properties; and more specifically, Fasssi Fehri (2018, Ch 5) for motivation of this terminology.

tive, rather than group (or collective), and its subject as a normal plural, just like the sound plural in (43):[14]

(43) *l-qaatil-uu-na   xtaba?-uu* (* *xtaba?-at*)
    the-killer-PL-NOM hid-PL   (* hid-UNIT)
    'The killers hid.'

(44) *l-qatal-at-u   xtaba?-uu*
    the-killer-FEM-NOM hid- PL
    'The killers hid.'

Hence depending on the 'perspective' of the speaker, the clause containing the plurative DP can have a collective/unity interpretation, as in (42), or a distributive/ scattered interpretation, as in (44). This core case of overt plurativity (which has no equivalent in English) is in fact limited to human plurals. It can also extend to some human group nouns (such as *naas* 'people'), and to human broken plurals, as illustrated in (45), and (46):[15]

(45) *n-naas-u   ġariib-at-un   (ġariib-uuna)*
    the-people-NOM strange-UNIT-NOM (strange-PL-NOM)
    'The people (as group) are strange.'

(46) *r-rijaal-u   mutaraddid-at-un   (mutaraddid-uuna)*
    the-men-NOM hesitant-UNIT-NOM (hesitant-PL-NOM)
    'The men (as group) are hesitant.'

In addition to the diversity of patterns of number agreement on verbal predicates (which are all arguably semantico-syntactically conditioned), there is also a variation on adjective predicates, although of non-equal importance. The most important one, I believe, is based on the human/non-human distinction, as exemplified in (47) and (48):

(47) *l-masaaf-aat-u   kabiir-at-un*
    the-distance-FEM.PL-NOM big-FEM-NOM
    'The distances are big.'

---

14 Such facts make the operation of Agree closer to unification, where both the probe and the goal contribute interpretable values. The discussion is worth pursuing, but it takes us too far afield. See e.g. Dowty & Jacobson (1988) for a very relevant discussion of this issue.
15 These considerations clearly indicate that the plurative is neither driven by the group property (since groups like *šaSb* 'people' or *saḥb* 'companions' cannot trigger plurative agreement), nor by broken plural morphology, since non-human broken plurals do not alternate as regular plural/ plurative (contrary to what has been proposed by Zabbal *ibid*, or Acquaviva 2008, among others).

(48) *l-fatay-aat-u*   *kabiir-aat-un*   *(*kabiir-at-un)*
      the-girl-FEM.PL-NOM big-FEM.PL-NOM (big-FEM-NOM)
      'The girls are big.'

The contrast indicates that (in the normal case) a non-human feminine plural does not trigger a feminine plural agreement on the adjective (but only feminine singular), unlike a human feminine plural which must do so.

   A second important distinction is the plural/plurative distinction, instantiated with examples (45) and (46) above. Furthermore, it is worth observing that the alternant to (47) where the adjective is a normal plural, is not ungrammatical, as shown in (49):

(49) *l-masaaf-aat-u*   *kabiir-aat-un*
      the-distance-FEM.PL-NOM big-FEM.PL-NOM
      'The distances are big.'

In this use of plural feminine agreement, the adjective phrase means something like 'very big distances'. The plural then functions as a *degree* (or evaluative) plural modifying the adjective, rather than a mere agreement marker. It requires a semantic treatment like other modifier plurals. As far as allomorphy is concerned, the sound plural in (48) or (49) can alternate in this context with broken forms of the adjective like *kibaar* 'big.pl', *ṭiwaal* 'large.pl', etc.[16]

3.2 Syntactic projections

A significant body of the literature converges on the idea that number is found at many levels or projections of structures, starting earlier with structuring the scale of (in)dividuation of entities or events, then quantities, measurements, event complexes, or temporal phrases. Due to the huge literature on the various ontologies and domains, I will limit myself here to sketch only some properties of Number projections. For the sake of concreteness, I will assume basically the architecture given by Acquaviva (20017a and b) for the nominal domain. Borer's DivP is split up into two distinct projections, AtomP and UnitP (Fassi Fehri 2018, Ch 5 for further motivation). Details about the event domain and its organization are provided in

---

16 Other differences are found in Ferguson (1959, 1989), Fassi Fehri (1988a, 2012, 2018). In particular, in alternating VSO and SVO orders in clauses, number is absent in VSO, and present in SVO for complex reasons (see Fassi Fehri 1993, Lowenstamm 2011; see also Shlonsky 1997, Benmamoun 2000). As for plurals on adjectives, it is important to observe that the variation is not free, and that the plural is not always an agreement plural, contrary to what Belnap (1999) suggests.

Section 4. Let us then start with a simple representation of the DP structure as in (50):[17]

(50)

```
        DP
         \
          QP
           \
            NmrP
             \
              NumP
              /  \
           Num   UnitP
                  /  \
               Unit   AtomP (nP)
                        \
                        √P
```
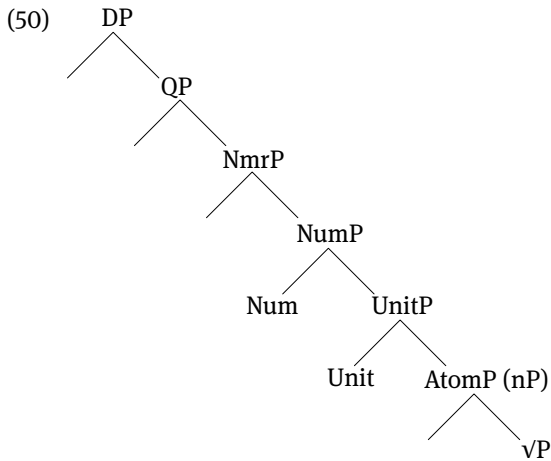
A nominal derivation starts with a root √. This root stage, modelled as a syntactic projection, √P, should not be devoid of any content, and would include a conceptual base (reflecting part of the content of a traditional Semitic root) which can serve for various pre-categorial (syntactic) derivations. It is then possible that an abstract (semantico-syntactic) form *KLB* (for 'dog') is pre-categorial, and that such a root denotes a concept DOG, which means that it is not totally free of any semantic specification. This form undergoes preliminary operations before it becomes *n*, *kalb*. Alternatively, a root is "just a purely differential index with no conceptual content" (Harley 2014, Acquaviva 2017a, b), and *n* stands for a noun concept, naming an entity type *e* (there being no pre-existing elements of content labelled by roots). A third option is that it is the root √ which identifies the basic entity type, and the categorized root [√]n is 'severing' it to a more specific meaning. As a corollary, nouns, verbs, and adjectives derive from the 'same root' (associated with some abstract form and meaning). Hence the root *KLB* with the general abstract sense KLB (i.e. whatever abstract 'dogness' means) serves as a base not only to name the ani-

---

17 DP is the locus of reference, deixis, definiteness, and can be split into DP and KP to provide a special location for case. The projection nP, headed by *n* (Marantz's 1997 categorizer, the existence of which is disputed by Borer 2005, 2013, Adger 2013) can be introduced over √P. Acquaviva (2017b) distinguishes an *n* functor, notated [√]ₙ, which names an entity type *e*, and is above √, from a higher PΣ functor, the property of sums, which introduces a variable, and creates a lattice. For lack of space, I will not discuss these elaborations in detail. NmrP and QP are placed here as separate projections for numeral phrases and quantifier phrases, just as is NumP (or #P, the locus of quantity Number). For more granular functional projections in the DP, as well as in the clause structure, see Rizzi and Cinque (2016).

mal *kalb*, which has a general and a particular nominal meaning DOG (as type or token), but also to construct the verb *kaliba*, literally 'dogged', meaning 'to be affected by a particular disease from dogs, or raging', or the adjective *kaliib*, literally 'doggy', but meaning 'affected with rabies, raging', or the more complex verb *takaalab* 'to be behaving inappropriately like a dog, to dogfight', or its deverbal metaphoric *takaalub* 'engaging in a nasty action', etc. Since there is no way to derive the common core meaning of these variously categorized words directly from the nominal sense (pointing to the entity, with its shape, integrity, etc.), then the common core sense of those words would be lost, at the cost of not adequately accounting for some traits of the speaker competence in detecting lexical relatedness.

What comes after the root is the nominal category, or the [√]n merger, then a higher DivP projection, as in Borer (2005), where 'dividing reference' and partitioning is operated, to make nPs numerable or countable. I assume that there are two stages or routes to division or partition: *atomicity*, projecting as atomP, and *unicity*, projecting as unitP. Nouns like *kalb* 'dog', *samak* 'fish', and *zayt* 'oil' all denote *kinds*, but the former pair denotes kind individuals that have more structure than the last noun denoting kind mass (see Carlson 1977 for more on kinds). It is reasonable to think of this individuality as an 'atomicity' of some sort (with concrete natural atomicity at the core), and identify it with an atomP projection in syntax, whereas masses lack such projection (because these are 'atomless'). I call unit or *unity* a necessary component in the structure of the singulative *samak-at* 'fish-UNIT', derived from the 'collective' kind *samak* 'fish', to become directly countable. If the nP projects unity, or unitP, then the right entity needed for counting is built, to achieve the right result. The idea is then that at a first stage both singulars project atomP, and at a second stage, unitP is projected, to form the unit *samak-at*. The two singulars *kalb* and *samak-at* are then not equal because they do not have the same derivational history, nor the same syntactic behaviour. Available accounts of countability (as far as I can tell) do not normally provide a mechanism to differentiate the two kinds of singulars, since they assume that both are born as masses, and continue to be masses until they are divided on equal bases (in DivP; see e.g. Chierchia 1998).

Thus, two features are central to the Number system I am proposing, as shown earlier: the *atom* feature and the *unit* feature. The basic idea is that number information is not just about atomicity; it is also about unity. Both contribute somehow to 'dividing reference' in the sense of Quine (1960), and intersect with Div in Borer (2005), although, my system is less 'coarse'. Not only are individuals born or built in the grammar as atoms (or as the bottom parts of 'division'), but they can also be built as units, or unities (possibly assembling individual atoms to form a unit or unity, or 'refining' a potential sum, to build a unit from it).

3.2.1 Syntax of head plurals: NumP and UnitP

The syntax of morphological number (associated with nouns or verbs) differs, depending on the semantics involved. One sort can be argued to be a *head*, an extend-

ed projection of e.g. the nominal head, and interpreted as such. It is generally assumed that the heads of the various functional projections in the extended domain are related by *selection* (e.g. Number selects a nP, just like T selects a vP, etc). From the semantic side, the value of number restricts the individuable variable provided for the interpretation of *n* as being 'one', 'two', or 'more than one'. I include in this class #P and UnitP plurals. Their positions in syntax, and with respect to each other, or to other projections, namely DivP, or singulative UnitP, is a matter of intense debate.

There is reason to think that the plural of the singulative is higher than the singulative UnitP (or DivP); see Fassi Fehri (2004, 2012, 2018) and Mathieu (2014) for detail.[18] As for the UnitP plural (or plurative), it differs from the normal #P in that it appears to fulfil two functions (simultaneously), instead of just one: it (often) mixes (a) 'summation' and (b) 'unitization' in a single morphological form. But the two functions are sometimes expressed by separate morphologies. For example, a broken plural expressing a sum plural (hence a #P or NumP) can be taken as a base to form a UnitP, by attaching a plurative morpheme to the plural noun (or nP), as in (51):

(51) a. *tilmiiḏ* → *talaamiiḏ*
　　　 student　student.PL student.PL-UNIT; students as group
　　 b. *barbar* → *baraaber* → *baraaber-at*
　　　 berber　 berber.PL; Berbers　 berber.PL-UNIT; Berbers as group

If so, then UnitP can be higher than NumP, suggesting that a plural can be made a unity (or UNIT) in syntax.

3.2.2 Syntax of modifer plurals

Another sort of plurals that cannot be seen as sums, because they are not heads, but rather modifiers of the noun phrase, include so-called double plurals, and plurals of masses. Consider first the double plural in (52):

(52) *daʕ-ka　 mina l-ʔaqaawiil-i　　 (l-ʔaqwaal-i)*
　　 leave-you　from　the-saying.PL.PL-GEN (the-saying.PL-GEN)
　　 'Forget about a lot of what is said' (literally: the 'many sayings').

Construction (52) does not strictly mean a (re-)pluralization of the plural, creating an upper sum from the initial sum. Rather, it qualifies a high amount or degree of

---

18 The alternative is that it is just a form of agreement (with a hidden numeral), as explained earlier.

plurality. The second plural can then be seen as a modifier plural, presumably an adjunct modifier (or degree) to the NumP headed by the plural. I assume that this kind of pluralization is syntactic, rather than lexical.

Consider now a case of mass plural, exemplified in (53):

(53) *saqaṭ-at ṯuluuj-un   kaṯiir-at-un*
       fell-FEM snow.PL-NOM many/lot-FEM-NOM
       'A lot of snow fell.'

In one reading of (53) at least, what is meant is a big amount of snow that has fallen, rather than 'more than one snow', or 'many snows'. In other contexts, 'snows' is coerced to be countable, and means 'kinds of snow', 'snows in many places', etc. The plural then is interpreted as a *#P*. But in the mass reading of (53), the plural corresponds to 'a lot', rather than 'many'. To obtain this reading, the plural is adjoined to the head noun.[19]

Another interesting case of non-head plurals discussed in the context of verbal number is the 'root' plural. It is important to clarify how it operates morphosyntactically, in view of its interpretative ingredients. First, note that the pluractional morpheme is part of the composition of the event (rather than the verb). The event base can be simple, amounting to a single event, or it becomes complex, when a pluractional 'intensive' is introduced. The complexity is similar to that of the causative, the reflexive, etc., where more than one event is involved. Second, it is reasonable to think that the complex event base is a complex root which is common to verbs, (event) nouns, or (adjectival) participles, although in different guises or configurations. Once the plural is taken as a part of the complex root, it is a root itself, as postulated in Fassi Fehri (2012, Ch 11). Indeed, Lowenstamm (2014) has more explicitly argued for a general view that so-called derivational affixes are roots rather than categorizers. Let us then call this plural a *root plural*. The latter should apply to roots that are events, *e*, not to roots that are objects, *o* (or 'things'; see Hale and Keyser 2002 on this view). When roots become categorized as verbs, verbal nouns, or verbal adjectives, they can no longer pluralize by any plural morpheme. Hence, there is no *v* plural or vP plural that has a parallel semantic content to *n* plural or # plural. At a phrasal level, v is associated with its arguments (or thematic/participant roles) in the VP or vP, to form new complex events, and the root plural

---

19 The agreement reflected here is similar to the pattern of agreement found in pluralia tantum constructions, as in the following Moroccan Arabic example:

(i) *mšati ktir-a*   (* ktir)
      rains many/lot-FEM   (* a lot)
      ('There is) a lot of rain.'

Thus although the plural *mšati* has no singular counterpart, it agrees in feminine gender, like a non-human plural.

can still operate on these participants, the event complexity being composed of the event proper and the participants. See Section 4 below.

4 Semantics and discourse

Plurality as a morpho-syntactic device (or 'inflection') expressing 'more than one' in plural nouns or verbs (in addition to adjectives), or 'more than two' (in the case of pronouns), needs a framework that can capture the common core of both object and event reference. In verbal configurations, plurality is expressed on the event, the participants, or the adjunct (adverbial) modifiers. My analysis builds on significant contributions by Link's (1983) conceptualization of algebraic atomic semantics, Krifka's (1992) treatment of counting event variables, Higginbotham's (1985, 2000) and Parsons' (1990, 2000) neo-Davidsonian semantics, as well as 'liberal' event extensions by Higginbotham (2005) and Ramchand (2007). The 'one generative engine' grammar (including both the syntax and the lexicon), as originally conceived by Hale and Keyser (2002), Marantz (1997, 2001), Borer (2005), and Harley (2014), among others, is used as an interface to construct meaning composition and event complexes.

4.1 Features and lattices

Central to the theory of number is to establish a set/inventory of empirically motivated semantic features that apply compositionally to nominal, pronominal, or verbal lattices, and to categories such as Num, n, Root, etc. These features are made available for natural number systems, and their natural classes. In Link's lattice model for number, the domain D_e of entities of type e contains both atoms (singularities) and their sums (pluralities). The model is structured as an atomic join semi-lattice.[20] The part-of relation of the mereology is ≤, and the join operation is ∪. As pointed out by Harbour (2014), if lattices are domains of variables, then number features are restrictors on these variables. Thus, given a lattice like (54), 'singular' restricts variables to range over atoms at the lower level, given in (55a), dual to the intermediate level (55b), and plural to the supremum (55c), and eventually medial (55b).[21]

---

20 A join semi-lattice (or upper) is a partially ordered set that has a *join* for any nonempty finite subset. A *meet* (or lower) semi-lattice is a partially ordered set which has a meet (or greatest lower bound) for any nonempty finite subset. A lattice which has both a join and a meet is full or complete. Join is represented in Champollion & Krifka (2016) by the symbol ⊕, and meet by the symbol ⊗. Only join is needed here.

21 As observed by Nouwen (2016), the desired structure is a *complete* atomic join semi-lattice. It is complete because the domain of entities is closed under ∪. It is *atomic* since all atomic parts of sums in the domain are part of the domain. It is a *semi*-lattice, since it uses only join, and no meet,

(54)

```
            {a + b + c}
          /      |      \
     {a + b}  {a + c}  {b + c}
        |        |        |
       {a}      {b}      {c}
```
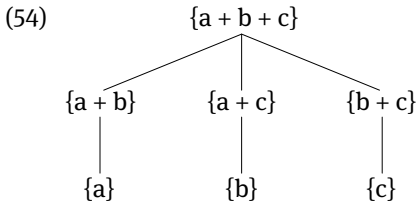
(55) a.   {a, b, c}
    b.   {a, b}   {a, c}   {b, c}
    c.   {a}, {b}, {c}
    d. $a \cup b = \{a, b\}$
    e. $\{a, b\} \leq \{a, b, c\}$
    f. $a \leq \{a, b\}$

In Fassi Fehri (2018, Ch 5), I assumed that the atom feature is bivalent, with three valuations: (a) [+ atom] for singulars, (b) [– atom] for plurals, and (c) [± atom] for general number. I take the denotation of a singular like *kalb* 'dog' to be associated with the bottom line of the lattice, the dual *kalb-aani* to denote medial sums, and the ('strong' or exclusive) plural *kilaab* to be the supremum of the lattice. More frequently, the plural subsumes the dual, or it may behave like general nouns, such as *samak* 'fish'. In this situation, the 'inclusive' plural denotes the complete set in the lattice, like a general noun (or the kind individual).[22]

The system proposed introduces a *unit* feature, [± unit], which differentiates the various singularities and pluralities along this specification. Thus *samak-at* 'fish-unit' is distinct from *rajul* 'man', as a singulative and singular, respectively, just as the plurative *qatal-at* 'killers (as a group)' in (43) is distinct from the plural *qaatil-uun* 'killers' in (44) above. Both the singulative and the plurative are [+ unit], while the singular and the plural are [– unit]. I assume that 'lexical' groups like *lajn-at* 'committee' or *ʔusr-at* 'family' are [± unit], hence accounting for their double behaviour, as described notably by Barker (1992), among others. Thus, some feature specifications of the more traditional divisions of singularities and pluralities and groups are given in (56) to (59):[23]

---

otherwise it would be a *full* lattice. In Champollion & Krifka (2016), a join semi-lattice is defined as a structure for a set *S* as a two-place operation on S called join, and symbolized by $\oplus$, such that for each two elements $x, y \in S$, the join $x \oplus y \in S$.

22 I concur on this point with Ojeda (1992), Rullman and You (2006), and Zabbal (2002). See also Sauerland (2003). If so, then one advantage of the bivalent feature system is to make available a straightforward account of the inclusive/exclusive divide of plurality. See Harbour (2011) for more motivation on bivalency.

23 For both Barker (1992) and Landman (1989), groups are taken as atoms, although they are both plurals and atoms for the former, and 'impure atoms' for the latter (derived via the group formation operator ↑). In my view, groups are not atoms, but unities. For detailed motivation, see Fassi Fehri (2018, Ch. 5), and earlier references given therein.

(56) a. *rajul* [+ atom, –unit]
    b. *rijaal* [– atom, –unit]

(57) a. *samak* [± atom, –unit]
    b. *samak-at* [± atom, +unit]

(58) *ʔusr-at* [– atom, ± unit]

(59) *qatal-at* [– atom, +unit]

The natural picture that emerges of the singular and plural classes and their cross-classification is then as follows (Sgv = singulative, Plv = plurative, G = group, K = kind individual):

(60) *Singular and plural classes*

	+ atom	*− atom*	*+ atom*
+ unit		**Plv**	**Sgv**
− unit	**Sg**	**Pl**	**K**
± unit		**G**	

Needless to say, more research is needed to refine this picture. Other features may also be necessary to enlarge the descriptive adequacy of the theory of natural Number systems.[24]

4.2 How can events be numbered

Let us turn now to events. Events have been central in the semantic understanding of pluractionality and distributivity on a par with plurality in the nominal domain (Schein 1993, Lasersohn 1995, Newman 1990, Landman 2000, among others). Links between nominal and temporal reference have been attempted in appealing work

---

[24] I have not included *mass* because it does not include any individuating features, according to most analyses, notably Borer (2005), although see Chierchia (1998) for a different analysis. Aggregates, as discussed by Acquaviva (2008), are not included here either. Harbour (2011, 2014) provides motivation for [± additive] and [± minimal] features, etc.

by various authors, including Bach (1986) and Verkuyl (1993). Krifka (1992) has explored the lattice model of Link (1983) to provide a unifying treatment of objects and events.[25]

The use of a Davidsonian event argument as an essential part of the denotation of verbs has become standard since at least Higginbotham (1985), and its motivated extension to participant roles by Parsons (1990) has also been generally accepted. Higginbotham also proposed that nouns are predicates which license e (x or i for individual) variables that need to be bound by a determiner (D) to be saturated, and their thematic role discharged (through theta-marking). Verbal predicates license e (eventuality) variables for the sake of event closure by Tense, or adverb modification (through theta identification). Thus, a sentence like (61) has a basic semantics as in (62), represented in (63) (Tense interpretation aside):

(61)   John walked rapidly.

(62)   There was a walk by John, and it was rapid (for a walk).

(63)   ($\exists$e) walked (e), & (John, e) & rapid (walk, e)

By theta identification, e in the adjective gets associated with e in the verb. Details of representation have changed since the introduction of participant roles via e in Parsons (1990), and more elaborations of complex event composition have been brought up by Higginbotham (2005) and Ramchand (2007), among others. Thus, the event variable is present not only in the description of 'verbs', but also in that of their arguments or adjuncts. The number of necessary e's then turns out to be many, and "apart from nominal arguments, verbs can be modified by an unpredictable number of linguistic phrases ... to be integrated into the description of the event" (Ramchand, *ibid*: 478). But as critically remarked by Maienborn (2005), the event position is extended to predicates of all sorts, resulting in an "infinite regress of eventualities". However, this liberal and multiple view is defended by Higginbotham (2005) and Ramchand (*ibid*), and I will assume a version of it for the sake of implementation, with the effect that multiple event variables in complex events are constantly modified in rich functional projections (as in Cinque 1999, Ernst 2002, Kratzer 2008, among others).

---

25 Krifka (*ibid*, 33) conjectures that "to handle the semantics of cumulative and quantized reference, we must provide for the semantic operation of joining two individuals to a new individual ... our model structure must be of the form of a lattice (cf. Link 1983). ... to cover event predicates ... two non-overlapping sorts of entities, objects (characterized by a predicate O), events (characterized by a predicate E), and times (characterized by a predicate T). The extensions of O, E and T have the structure of a complete join semi-lattice without a bottom element ... a function T from the extension of τ to the extension of T, the temporal trace function; this function maps an event to its 'run time', or temporal trace". The join of the temporal traces of two events equals the temporal trace of the join of these events.

4.3 Counting event units and counting times

Let us take examples of the so-called 'verbal' plurality, to make precise what is the contribution of each element composing the complex event structure, and its numberability or countability. Assume activity verbs are cumulative, or non-quantized, as in Krifka (1992), or more precisely *kinds*. The pluractional marker can then be a special plural (in the sense of Souckova *ibid*), a modifier, or more precisely a root plural.[26] A root plural composes with its base root to produce a new complex root, more like any modifier does in the root of manner verbs. Verbs like *ḥaddaqa* 'to stare' and *ḥamlaqa* 'to goggle' have a manner component in their complex event lexical structure, before they get categorized. Basic atelic verbs become quantized or telic when they are constructed with incremental objects. If thematic roles are event roles, as in Parsons (*ibid*), then an accomplishment predicate like (64) can be represented as in (65), accounting for the fact that an internal event composition with a thematic role may result in a 'countable' verb phrase:

(64) *Ɂakala r-rajul-u   tuffaaḥ-at-an   marrat-ayni*
      ate   the-man-NOM apple-UNIT-ACC time-DUAL
      'The man has eaten an apple two times.'

(65) λe′Ǝe [*Ɂakal* (e) & Agent (*r-rajul*, e) & Theme (*tuffaaḥat*, e)]
      & Time (e′, e, *marrat*) & n (time) = 2

In one interpretation of (64), what is counted is the complex accomplishment event, 'eating an apple', rather than the act of 'eating'. That interpretation may result in the man eating two apples, rather than just one.

   Cognate event units (in contrast to kind events) can serve to count the number of the individual units of the event taking place, as in (27) to (29) above. The event unit is perceived as a verbal classifier, because it makes the event countable. But since it involves the event, rather than the verb itself, it is more appropriate to call it an *event classifier*. The function of the classified cognate is similar to that of an adverbial adjunct, rather than to a participant object. Cognates can be attached as adjuncts to a verb phrase already containing an object theme, as (66) illustrates:

(66) *Ɂakala r-rajul-u   tuffaaḥ-at-an   Ɂakl-an   taamm-an*
      ate   the-man-NOM apple-UNIT-ACC eating-ACC complete-ACC
      'The man has eaten an apple completely.'

---

26 See Wiltschko (2008) and Butler (2012) for discussion of properties of modifier plurals.

This suggests that the adjunct cognate is attached and interpreted higher than the internal object, and the structure of the event composition is roughly as in (67); note that the cognate event reduplicates the verbal event:[27]

(67) λe′∃e [ʔakal (e) & Agent (*r-rajul, e*) & Theme (*tuffaaḥat, e*)] & (*taamm*, e′)
       & (*ʔakl*, e, e′)

It is suggestive that the verbal predicate (unlike the nominal) is compatible with two countings. This is the case where the 'verb classifier' is used in the context of counting the number of times that the event takes place (as postulated by Krifka 2013; see also Zhang 2012). Then the event is counted twice, as in (68):

(68) *raqaṣa r-rajul-u   raqṣ-at-ayni   t̲alaat̲-a   marr-aat-in*
      danced the-man-NOM dance-UNIT-DUAL three-ACC time-PL.FEM-GEN
      'The man danced two dances three times.'

(69) λe″∃e′, e [*raqaṣ* (e) & Agent (r-rajul, e) & Event (e′, *raqṣ*) & n (e′) = 2
       & Time (e″, e′, *marrat*) & n (time) = 3

The temporal role involved in counting time occurrences or units can be seen as of type *n* (for number), as proposed by Scha 1981, Krifka 1992, Zabbal 2002, Fassi Fehri 2018, among others. Likewise, the temporal role in counting events can also be of type *n* (*ibid*).

Due to lack of space, I point only very briefly to instances of plurality that are discourse driven. One example described, depending on the viewpoint of the speaker, is whether the plural is perspectivized as a 'unity' (or collective), or as a non-unity (or distributive). This is represented by the plurative/regular plural alternation (as in Leiss's 1994 'perspectivization' view). A second example is when non-human nPs are personified and used with a regular plural agreement.

5 Conclusion

I have provided a broad view of the Arabic number system, with ingredients of a general theory of number, and with special attention to the specifics of nominal, verbal, and pronominal numbers. It turns out that the nominal number subsystem manifests the richest variation in terms of functions and exponence, although the other subsystems can be equally treated in parallel, at least tentatively, with lan-

---

27 See Jones (1988) for an early treatment of cognate objects in syntax, and Fassi Fehri (1988b) for Arabic.

guage specific choices. Various new number and individuation subtypes have emerged, including root plural, n plural, num plural, event classifier, plurative, singulative, head plural, and modifier plural, among others. Number and individuation classes have also been motivated. Needless to say, more systematic comparisons of seemingly distinct languages are needed, in addition to a more integrative theory.

Acknowledgements

Abbreviations

ACC	accusative
CL	classifier; CLP classifier phrase
D	determiner; DP determiner phrase
DIVP	divider phrase
DL	dual
e	event
FEM	feminine
G	group
GEN	genitive
INTENS	intensive
K	kind
KP	case phrase
MASC	masculine
n	noun; number argument
NMR	numeral
NOM	nominative
nP	noun phrase
NUM	number
PERS	person
PL	plural; PLV plurative
QP	quantifier phrase
#P	quantity phrase
√ root;	√P root phrase
SG	singular
SGV	singulative
UNIT	unity, unitizer

References

Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford University Press.

Acquaviva, Paolo. 2017a. Number in Language. *Oxford Research Encyclopedia in Linguistics*. 56 pp. https://doi.org/10.1093/acrefore/9780199384655.013.61

Acquaviva, Paolo. 2017b. Categorization as noun construction. Gender, number, and entity types. https://ling.auf.net/lingbuzz/003317

Adger, David. 2013. *A Syntax of Substance*. Cambridge, MA: MIT Press.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16.

Barker, Chris. 1992. Group Terms in English: Representing Groups as Atoms. *Journal of Semantics* 9. 69–93.

Belnap, Kirk. 1999. A New Perspective on the History of Arabic Variation in Marking Agreement with Plural Heads. *Folia Linguistica* 33. 169–186.

Benmamoun, Elabbas. 2000. *The Feature Structure of Functional Categories*. Oxford: Oxford University Press.

Boas, Franz. 1911. Introduction. In Franz Boas (ed.), *Handbook of American Indian Languages*, 1–83. Washington: Smithsonian Institution.

Borer, Hagit. 2005. *Structuring sense. In Name only*. Oxford: Oxford University Press.

Borer, Hagit. 2013. *Structuring Sense. Taking Form*. Oxford: Oxford University Press.

Butler, Lindsay. 2012. Cross-linguistic and experimental evidence for non-number plurals. *Linguistic Variation* 12(1). 27–56.

Cabredo Hofherr, Patricia & Brenda Laca (eds.). 2012. *Verbal plurality and distributivity*. Berlin: De Gruyter.

Carlson, Gregory. 1977. Reference to kinds in English. PhD thesis. University of Mass at Amherst.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6. 339–405.

Champollion, Lucas & Manfred Krifka 2016. Mereology. In Maria Aloni & Paul Dekker (eds.), *The Cambridge Handbook of Formal Linguistics*, 369–388. Cambridge: Cambridge University Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Cusic, David. 1981. Verbal plurality and aspect. PhD dissertation. Stanford University.

Doetjes, Jenny. 2008. Counting and degree modification. *Recherches Linguistiques de Vincennes* 37. 139–160.

Dowty, David. 1979. *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Reidel.

Dowty, David & Pauline Jacobson. 1988. Agreement as a semantic phenomenon. *ESCOL* 5. 95–108.

Durie, Mark. 1986. The grammaticization of number as a verbal category. In *Proceedings of BLS* 12, 355–370. Berkeley: Berkeley Linguistics Society.

Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Fassi Fehri, Abdelkader. 1988a. Agreement, Binding, and Coherence. In Michael Barlow & Charles Ferguson (eds.), *Agreement in Natural Language*, 107–158. Stanford: CSLI.

Fassi Fehri, Abdelkader. 1988b. Cognate Objects as Arguments. Ms. Cambridge: MIT & Rabat: Mohammed V University.

Fassi Fehri, Abdelkader. 1996. Distributing features and affixes in Arabic subject verb agreement paradigms. *Linguistic Research* 1(2). 1–30. Rabat: IERA.

Fassi Fehri, Abdelkader. 2001. Causativity, transitivity, and iterativity as pluralities. *Linguistic Research* 6(1). 7–80. Rabat: IERA.

Fassi Fehri, Abdelkader. 2004. Nominal classes and parameters across interfaces and levels, with a particular reference to Arabic. *Linguistic Variation Yearbook* 4. 41–108. Amsterdam: John Benjamins.

Fassi Fehri, Abdelkader. 2012. *Key Features and Parameters in Arabic Grammar*. Amsterdam: John Benjamins.

Fassi Fehri, Abdelkader. 2013. The new Arabic lexicon and in its features. Talk presented at the Linguistic Colloquium of the University of Tunis Manounba. April 10.

Fassi Fehri, Abdelkader. 2018. *Constructing the Feminine to Mean. Gender, Number, Numeral, and Quantifier extensions in Arabic*. New York: Lexington Books.

Ferguson, Charles. 1959. The Arabic koine. *Language* 35. 616–630.

Ferguson, Charles. 1989. Grammatical Agreement in Classical Arabic and the Modern Dialects. *Al-ʕ-Arabiyya* 22. 5–17.

Ferrando, Ignacio. 2006. The plural of paucity in Arabic and its actual scope. In Sami Boudelaa (ed.), *Perspectives on Arabic linguistics* XVI, 39–62. Amsterdam: John Benjamins.

Glanville, Peter. 2018. *The Lexical Semantics of the Arabic Verb*. Oxford: Oxford University Press.

Greenberg, Joseph. 1991. The Semitic intensive as verbal plurality. In Alan Kaye (ed.), *Semitic Studies in Honour of Wolf Leslau II*, 577–58. Wiesbaden: Harassowitz.

Hale, Kenneth & Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge MA: The MIT Press.

Harbour, Daniel. 2011. Descriptive and explanatory markedness. *Morphology* 21. 223–245.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90. 185–229.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40(3–4). 225–276.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16. 547–93.

Higginbotham, James. 2005. Event positions: Suppression and emergence. *Theoretical Linguistics* 31. 349–58.

Jones, Michael. 1988. Cognate Objects and the Case-Filter. *Journal of Linguistics* 24(1). 89–110.

Kratzer, Angelika. 2008. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event Structures in Linguistic Form and Interpretation*, 269–300. Berlin: Mouton de Gruyter.

Krifka, Manfred. 1992. Thematic relations and links between nominal reference and temporal constitution. In Ivan Sag & Anna Szabolcsi (eds.), *Lexical matters*, 29–53, Stanford: CSLI.

Krifka, Manfred. 2013. Measuring and Counting in the Nominal and the Verbal Domain. *Workshop on Countability*. Universität of Düsseldorf.

Landman, Fred. 1989. Groups, I. *Linguistics and Philosophy* 12(5). 559–605.

Landman, Fred. 2000. *Events and plurality*. Dordrecht: Kluwer.

Lasersohn, Peter. 1995. *Plurality, conjunction and events*. Dordrecht: Kluwer.

Leiss, Elisabeth. 1994. Genus und Sexus, Kritische Anmerkungen zur Sexualisierung von Grammatik. *Linguistische Berichte* 152. 281–300.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical Approach. In Rainer Bauerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, Use, and Interpretation of Language*, 303–323. Berlin: Mouton de Gruyter.

Lowenstamm, Jean. 2011. The Phonological Pattern of phi-features in the Perfective Paradigm of Moroccan Arabic. https://www.researchgate.net/publication/228836928.

Lowenstamm, Jean. 2014. Affixes as Roots. In Artemis, Alexiadou, Hagit Borer & Florian Schäfer (eds.), *The syntax of roots and the roots of syntax*, 230–257. Oxford: Oxford University Press.

Maienborn, Claudia. 2005. On the limits of the davidsonian approach. *Theoretical Linguistics* 31. 275–317.

Marantz, Alec. 1997. No escape from syntax. *UWPL* 4. 201–225.

Marantz, Alec. 2001. Phases and Words. Ms. New York University.

Mathieu, Eric. 2013. The plural of the singulative. *McGill Working Papers in Linguistics* 23. 1–12.

Mathieu, Eric. 2014. Many a plural. In Ana Aguilar-Guevara, Bert Le Bruyn & Joost Zwarts (eds.), *Weak Referentiality*, 157–182. Amsterdam: John Benjamins.

McCarthy, John & Alan Prince. 1990. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8. 209–283.

Meirav, Ariel 2003. *Wholes, Sums and Unities*. Dordrecht: Kluwer.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology*, 211–234. San Diego: Academic Press.

Nouwen, Rick. 2016. Plurality. In Maria Aloni & Paul Dekker (eds.), *The Cambridge Handbook of Formal Linguistics*, 267–284. Cambridge: Cambridge University Press.

Newman, Paul. 1990. *Nominal and Verbal Plurality in Chadic*. Dordrecht: Foris.

Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.

Ojeda, Almerindo. 1992. The semantics of number in Arabic. In Chris Barker & David Dowty (eds.), *SALT* II, 303–325. Columbus: Ohio State University.

Parsons, Terence. 1990. *Events in the semantics of English: a study in subatomic semantics*. Cambridge, MA: The MIT Press.

Quine, Willard van Orman. 1960. *Word and object*. Cambridge, MA: The MIT Press.

Ramchand, Gillian. 2007. Events in Syntax: Modification and Predication. *Language and and Linguistics Compass* 1(5). 476–497.

Rizzi, Luigi & Guglielmo Cinque. 2016. Functional Categories and Syntactic Theory. *Annual Review of Linguistics* 2. 139–63.

Rullmann, Hotze & Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In Klaus von Heusinger & Ken Turner (eds.), *Where Semantics Meets Pragmatics*, 175–196. Amsterdam: Elsevier.

Sauerland, Uli. 2003. A new semantics for number. In Rob Young & Yuping Zou (eds.), *SALT* XIII, 258–275. Ithaca, New York: CLC Publications.

Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal methods in the study of language,* part 2, 483–512. Amsterdam: Mathematisch Centrum.

Schein, Barry. 1993. *Plurals and events*. Cambridge, MA: The MIT Press.

Shlonsky, Ur. 1997. *Clause Structure and Word Order in Hebrew and Arabic*. Oxford: Oxford University Press.

Součková, Kateřina. 2011. *Pluractionality in Hausa*. Utrecht: LOT Publications.

van Geenhoven, Veerle. 2004. *For*-Adverbials, Frequentative Aspect, and Pluractionality. *Natural Language Semantics* 12. 135–190.

Verkuyl, Henk. 1993. *A theory of aspectuality: the interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press.

Veselinova, Ljuba. 2013. Verbal Number and Suppletion. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*. Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Wiese, Heike. 2012. Collectives in the intersection of mass and count nouns: A cross-linguistic account. In Diane Massam (ed.), *Count and Mass Across Languages*, 54–74. Oxford: Oxford University Press.

Wiltschko, Martina. 2008. The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory* 26. 639–694.

Wood, Esther. 2007. The Semantic Typology of Pluractionality. Ph.D. Berkeley: University of California.

Zabbal, Youri. 2002. The semantics of number in the Arabic noun phrase. MA thesis. University of Calgary.

Xavier Bach
6 Number in Occitan

Abstract: The marking of number in Occitan varies enormously according to dialect: most of western varieties mark it segmentally, while other use a mixture of vowel raising, stress shift and vowel length. A number of varieties also present asymmetries in the marking of number within the NP. Suppletion is only attested for number for 1SG and 2SG pronouns. There is no verbal number but subject (and sometimes object) number agreement on verbs. The second person plural of the verb is used for politeness address, but other agreement targets remain in the singular when referring to only one person. Some varieties present a restricted form of dual marking (elsewhere marked with a classifier) for items conceived as appearing in pairs. A number of lexical plurals can be recognised. One word, *monde* 'people', is used in the singular but triggers plural semantic agreement.

1 Overview

Occitan is a Romance language (Indo-European) spoken in the southern third of France, as well as in the Val d'Aran in Spain, in the village of La Guardia Piemontese in Southern Italy, and in sixteen valleys of the Piedmont in Italy (Sauzet & Oliviéri 2016:319). It is more properly defined as a dialect continuum which exhibits wide variability between the different dialects, as well as common features. Six main dialects can be distinguished: Gascon (sometimes referred to as a separate language), Languedocian, Limousin, Auvergnat, Provençal and Alpin. Even inside dialects there is wide variability between local varieties of the language. The language is well documented historically, with earlier records dating back to the eleventh century. Modern varieties tend to be more scarcely documented, except for a few of them (e.g. Quint 1996, 1998; Amaro-Péguy 2014; Sibille 2015). Phonology and the lexicon are the most well-known areas, with less emphasis put on syntax and morphology, and very little having been done on semantics and pragmatics.

The language has the typical inflectional profile of Romance languages, marking a number of features inflectionally and synthetically. Verbs mark tense, mood, person and number (of the subject), nouns mark only number, and determiners and adjectives mark both number and gender cumulatively. All nouns in the language have gender, opposing two values, masculine and feminine, which are only semantically defined for a subset of animate nouns. For other nouns, gender is lexically defined.

Number marking is obligatory on determiners for all dialects, on nouns and adjectives for most dialects (with the exception of Provençal and some northern varieties). Agreement in person and number with the subject is obligatory for all

verbs. All varieties distinguish two values for number, opposing singular and plural. Some varieties (Northern Lot and Dordogne) may be said to also present limited dual marking for items normally occurring in pairs, the distinction in number being made only by the presence of a special determiner.

The exponence of number in Occitan varies in a dramatic fashion in between varieties, and a large number of exponence relations are attested in some of the varieties. This is also the aspect of number that has been the most studied until now (Sibille 2009, 2011, 2016, Sauzet 2011, Lieutard 2004, Jagueneau 1979, Floricic 2010, McKenzie 2010, Mok 2008). Depending on the varieties, number can be marked segmentally (with the adjunction of a final -s segment or one of its allomorphs), with vowel alternations, vowel lengthening, accentual shift, or even for one variety through the use of tone (Sauzet 2011). Non-segmental marking is most common in Northern Occitan varieties (Limousin in particular).

There are no major differences between nominal and pronominal number, both marking the same set of values. The only difference is that number marking in pronominal paradigms is often achieved through suppletion rather than by simply adding a plural marker to a singular form. There is no major lexical or semantic split between nouns, nor are there many doublets for forming the plural. In addition, there is no evidence for verbal number in the language.

There is both NP-internal and clausal agreement implicating number. In most varieties, agreement for number is obligatory on all members of the NP (determiners, demonstratives, adjectives). Interestingly though, a number of varieties show different forms for the agreement of adjectives depending on whether they are prenominal or postnominal. In some varieties such as Provençal, the difference is about marking agreement or not: all prenominal elements agree in number and gender with the noun, while postnominal ones do not agree in number. In other varieties, there are simply different forms for prenominal and postnominal adjectives. This is an interesting phenomenon, which poses some theoretical challenges as to how it should be modelled. At the clausal level, there is agreement in number between the subject and the verb. Predicative adjectives also agree with the subject of the clause, as do past participles in periphrastic verbal forms with auxiliary *èsser* 'be'. In some cases, there is also agreement between the object and the non-finite form of the verb (past participle agreement on periphrastic forms including auxiliary *aver* 'have').

In Old Occitan, where there was still some remnants of a case marking system for some of the nouns (masculines and imparisyllabics) the segment -s was used sometimes in the singular (subject case), sometimes in the plural (non-subject case). But by the time of Late Mediaeval Occitan (15th century), the marking of number had been largely assumed by a final -s segment on all members of the NP.[1] Most of

---

1 The only exception is those prenominal markers (determiners, quantifiers, adjectives) which continue Latin second declension masculine plurals in -i, which are still preserved in some Pyrenean and Alpine valleys.

the actual present day variation between varieties can be assumed to have arisen from the differential treatment of final -*s*, which has a tendency to deletion combined with various compensatory phenomena (lengthening or raising of the final vowel, accent shift, development of tone).

Because the exponence of number varies dramatically in the language, I will present multidialectal data for the marking of number. In the rest of the paper, and because other aspects of number have been little studied until now, I will centre on a single Languedocian variety, spoken in Venés (Tarn),[2] but it is highly probable that aspects of the syntax and semantics of number in this variety can also apply to others. The language has a standard orthography designed to minimise the differences between dialects, but this orthography tends to obscure most of the relevant facts for number marking (consistent marking through final -*s*, although it may not be realized on some words, or may instead signal accent shift or another phenomenon). Thus examples are presented in a broad phonetic transcription instead of the standard orthography.

2 Pronominal, nominal, and verbal number

2.1 Generalities

Nouns and pronouns mark the same range of values for number (singular and plural), the main difference between them being that some pronouns mark number suppletively, while this does not occur for nouns. In addition, third person pronouns also mark gender, which is not the case for nouns (pronouns are in this respect closer to the marking exhibited by determiners and adjectives). There is no verbal number. Sauzet (2012) provides a useful typology of number marking in Occitan varieties, distinguishing two parameters: first whether the marking of number is 'sigmatic' (realized through the adjunction of a final -*s* segment or one of its allomorphs) or not, and second whether the marking of number is distributed across all members of the NP, or limited to its leftmost elements (determiners only, or all prenominal modifiers). This distinguishes four types of varieties with respect to the marking of number, and all four types are attested.

2.2 Pronominal number

There is a dichotomy for strong pronouns (subject and non-subject all have the same forms) between third person and non-third person pronouns: third person

---

2 I thank my main informant, Roger Aussenac, from Venés (Tarn), for his help in preparing this work. Except otherwise indicated, all examples come from this variety.

pronouns mark number segmentally in Venés, while non-third person pronouns mark number suppletively, as shown in Table 1.[3] Although second person plural is used in address to mark politeness, a specific subject pronoun is used in such cases, *bus*.

Tab. 1: Subject personal pronouns in Venés.

	SG	PL
1	jew	nuzawtres
2	ty	buzawtres
3	el / elɔ	eles / elɔs

Subject personal pronouns are never clitics, as Occitan is a pro-drop language (except for some of the Northern Occitan varieties). Direct object and indirect object pronouns are clitics, and present a different form, as shown in Table 2 (direct object) and 3 (indirect object). The bundling of grammatical features into forms is the same for direct objects and for subject pronouns, but indirect object pronouns are all pluralized through suppletion and do not mark gender. Non-third person pronouns also exhibit syncretism between their direct object and indirect object forms.

Tab. 2: Direct object clitic pronouns in Venés.

	SG	PL
1	me	nus
2	te	bus
3	lu / la / ba	lus / las

Tab. 3: Indirect object clitic pronoun in Venés.

	SG	PL
1	me	nus
2	te	bus
3	je	jur

---

3 For third person pronouns, I give first the masculine form, then the feminine form, as third person personal pronouns also mark gender. For object pronouns, there is an additional third person singular form which is referred to as a 'neuter' object pronoun, used when referring to a clause or to an indefinite thing. This form only exists in the singular, and triggers singular agreement on the past participle when used with auxiliary *aver* 'have'.

Some interrogative pronouns present a plural form (as well as a feminine form), such as *kyn* 'which' (M.SG *kyn*, F.SG *kyntɔ*, M.PL *kyntes*, F.PL *kyntɔs*), or *kan(t)* 'how many' but most interrogative pronouns and adverbs only have a singular, non gender specific form when they are used pronominally, such as *kal* 'who', *un(t)* 'where' and *kurɔ* 'when'. Demonstrative pronouns distinguish two degrees of distance from the speaker *akeste/ akestɔ, akestes, akestɔs* for proximal, and *akel / akelɔ / akeles / akelɔs* for distal. Reflexive pronouns mark number through suppletion for non-third persons, using identical forms to the direct object pronoun, but the third person reflexive pronoun takes the form *se* and does not distinguish number.

2.3 Nominal number

Nominal number is marked inflectionally, except for those varieties which do not mark number on nouns (Provençal, where prenominal determiners and modifiers are the only elements of the NP that express number). In some varieties, there are a number of inflectional classes depending on the final segment of the word in the singular, as well as its stress profile (see e.g. Usseaux presented below, Amaro-Péguy 2009). The animacy hierarchy does not seem to be relevant in the marking of number on nouns for any of the varieties. There is a small group of varieties of Occitan which mark a dual value for number for those items which are conceived of as occurring in pairs (trousers, pliers, etc), through the use of specific determiners combined with plural marking on the noun (see below). There does not seem to be classifiers in the language, other than quantifiers used to express specific instances or quantities of mass nouns, and the classifier [*ym parel de*] 'a pair of' used in those varieties which do not have dual marking to refer to those items that are conceived of as occurring in pairs (trousers, scissors, etc).

The exponence of number in Occitan varieties shows great variation (see e.g. Floricic 2010). Following the typology proposed by Sauzet (2012), I present examples of sigmatic and non-sigmatic marking of number in this section devoted to the exponence of number in Occitan. I present five different systems, which give a small idea of the amount of variation for the exponence of number in different varieties of the language.

The system present in Venés can be said to be sigmatic and distributed. Within the NP, determiners, some quantifiers, adjectives, and nouns are marked for number with one of the allomorphs of final -*s* (though see below about prenominal adjectives). Allomorphs are /s/ in front of voiceless stops, /-j/ in front of other consonants, and /-z/ in front of vowels. Words ending in a sibilant or the affricate /ts/ mark plural with the adjunction of /-es/. There are some irregularities, either because some nouns mark plural twice with /-s-es/, or for some determiners appearing in front of a vowel for which the double marking takes the form of /-j-z/, as in *da-j-z* 'of.the.M.PL'. This system is common across all Languedocian and Gascon dialects,

with variants in the type of allomorphs. Variations in allomorphy in Languedocian varieties is analysed in MacKenzie (2010), Bach (2012) and Jagueneau (1979).

Limousin varieties show an example of a non-sigmatic but distributed system: most members of the NP are marked for number, but through non-segmental means. The general rule can be expressed as a rule of final vowel lengthening, which generally also triggers a stress shift and the tensing of the vowel (Javanaud 1981:102). This initial lengthening can be said to be a compensatory rule following the earlier loss of final -*s* (and of final consonants in the dialect more generally). Those words which originally presented a final consonant do not vary, whether they present a long or a short final vowel, still retain a final -*r*, or end in a nasal vowel originally arising from a sequence V+N. On the contrary, words with paroxytonic stress and originally ending in a vowel show lengthening and tensing of the final vowel (for middle vowels only, high vowels remain identical), plus a stress shift to the final, long vowel (Javanaud 1981:101). Thus sg. *'awbre* pl. *aw'brej* 'trees', sg. *'vatso* pl. *va'tsa:* 'cows', sg. *'muli* pl. *mu'li:* 'mills', sg. *lu* pl. *lu:* 'wolves'.

There is one very interesting case in the Limousin of a single variety that has developed a prosodic contrast between the singular and the plural (Sauzet 2011). In the village of Saint-Julien-de-Crempse, recordings from the 1970's show a clear contrast associating a low pitch to the plural, but not with the singular of words. This development is unique among Occitan varieties. Sauzet analyses it as a development from a final -*s* to a vocalic marking -*w*, which triggered a lowering of the pitch accent on plural words.

Some varieties show highly complex systems of allomorphy for the plural marker based on the final segment of the word itself, forming phonologically constrained inflectional classes. This is in particular the case in the Alps, in Usseaux (Amaro-Péguy 2014). The realisation of the plural on nouns in Usseaux depends on two factors: the final segment of the word in its singular form, and the stress profile of the word (nouns, adjectives and participles). Paroxytonic words ending in -*e* in the singular change that vowel in the plural to -*i* for masculine nouns and to -*a* for feminine nouns. These words show a neutralization of gender in the singular, with a gender value correlated to inflectional class only visible in the plural. Other paroxytonic words are invariable. Oxytonic words ending in -*a* replace that vowel by -*o* in the plural, other oxytonic words ending in a vowel are invariable. Finally, for oxytonic words ending in a consonant, there are a number of patterns: words ending in -*c* or -*p* are invariable, -*el*/-*ol* is replaced by -*iaus*/ -*aus*, -*t* is replaced by -*s*, and words ending in -*r*, -*e*, or a nasal vowel simply add a segment -*s* to the singular form (Amaro-Péguy 2014:176). Past participles present other forms. This variety thus presents inflectional classes whose allomorphy is highly constrained by the phonology of the singular form of the word.

Provençal varieties present a non-sigmatic and non-distributed system, where only prenominal elements of the NP are marked for number. Nouns and postnominal adjectives remain unmarked. Prenominal elements mark number through a vo-

calic change of the final vowel of the determiner to *-i* or *-ɛj* depending on the varieties, as shown in (1):

(1) *li   pi'ʧun-i   'pumɔ sun   pu'lido*
     the.PL small-FPL apple be.PRS.3PL beautiful.F
     'The small apples are nice'

There is no evidence for a plural word, or for number suppletion on nouns in the language. In some varieties there are attestations of double plural marking with a first marker *-s* followed by the marker attested after a sibillant *-es*. In Venés, this is the case for the word for 'hair' *pɛl*, plural *pɛlses* instead of the expected *pɛls, and for a few other lexically determined items. Other varieties, in particular North Languedocian varieties, such marking has been extended to a larger number of masculine plural nouns, and also to third person plural masculine pronouns and determiners: *ɔkelses* 'these.M.PL', *elses* 'they.M.PL'. The interpretation of such nouns is not different from standard plurals. It seems that such doubling of the plural marker is more common for those varieties where final -s tends to be lost (*ɔkelse, else*), as is the case in northern Languedoc.

    There seems to be a restricted form of dual marking for varieties in Northern Languedoc and Dordogne, for twin-structured objects and for items that usually come in pairs (trousers, scissors, shears, etc; Bach 2012). This corresponds to what Buridant (2000) calls 'internal plurals' for old French. In fact, the plural form of numeral one *ynes* / *ynɔs* 'one.PL' is used for a single item in the indefinite, while *de las* is used for more items in the indefinite. The definite plural will always refer to a single item, thus creating a defective paradigm (no definite plural reading is possible). This usage seems to be relatively widespread for paired items of clothing, tools, and glasses.

(2) *la-s   te'naʎɔ-s*
     the.F-PL plier-PL
     'The pliers'

(3) *'yn-ɔ-s   te'naʎɔ-s*
     one-F-PL plier-PL
     'Pliers (one item)'

(4) *de   la-s   te'naʎɔ-s*
     PRTV the.F-PL plier-PL
     'pliers (more than one item)'
     (Blars, northern Languedocian, Bach 2012)

Such marking seems to have been common in Provençal until the 19th century (Ronjat 197:130–131). In other varieties, such items are marked with a classifier *ym parel*

de 'a pair of' which functions as a singulative. There does not seem to be a widespread use of classifiers outside this example, and to count quantities related to a mass noun.

Regarding the interpretation of the plural, pluralized mass nouns tend to elicit a sort reading. Thus *de βis* 'wines' will be understood as 'various types of wine'. Family names, with a definite plural determiner (the name does not take a plural marker) have a collective reading: *lus Awsenak* 'the Aussenac family'. There can also be a non-strict reading of the plural, with a possible interpretation of cardinality one, with questions only, as shown in example 5 (negative contexts will not elicit this reading, contrary to English, because of the presence of a negative quantifier, which triggers either singular or plural marking on the noun, example 6). When the quantifier is not present, negation always triggers plural agreement with indefinites (Occitan, like French, marks indefinite plural with a partitive with plural agreement on the noun (example 7), while true partitives for mass nouns trigger singular agreement on the noun, example 8).

(5)   - *Esk a'βεn*   *de*   *'pumɔ-s?*
      Q   have.PRS.1PL PRTV apple-PL
    - *n*   *a'βεm*   *pas k* *'ynɔ*
      of.it have.PRS.1PL not that one.F.SG
    '-Do we have apples? - We only have one'

(6)   *a'βεn*   *pas 'kad⁴ de*   *'pumɔ / 'pumɔ-s*
    have.1PL not none PRTV apple / apple-PL
    'We don't have any apples'

(7)   *a'βεn*   *'paj de*   *'pumɔ-s*
    have.1PL not PRTV apple-PL
    'We don't have apples'

(8)   *'baʎɔ*   *me*   *d*   *'ajɔ*
    give.IMP.2SG 1SG.DAT PRTV water
    'Give me some water'

Generic readings generally require the plural, as shown in example 9.

(9)   *lu-s*   *'kɔ-s*   *su*   *'braβe-z en ʦene'ral*
    the.M-PL dog-PL be.PRS.3PL nice-PL in general
    'Dogs are generally nice'

---

4 This word is a grammaticalized form of *kap* 'head'.

Pluralia tantum are commonplace, and always trigger plural agreement at both the NP and the clausal level. Similarly, singularia tantum always trigger singular agreement. Pluralia tantum tend to be more common for some types of meanings than others. They are particularly prevalent for names of diseases (*gawtisɔs* 'mumps'), some plants (*agyʎɔs* 'shepherd's needle'), meteorological events (*kastɛls* 'storm clouds'), internal parts of the body (*lɛws* 'lungs (of a pig or large animal)') special dishes (*farinɔs* 'gruel'), tools, games (*paʎetɔs* 'jonchets, kind of game'), and seasonal activities (*bendemjɔs* 'season of grapes harvesting') (Bach 2012:47–49 and 86–87). Most attestations are in fact lexical plurals presenting a different meaning than the attested singular form (*agyʎɔs* 'shepherd's needle', and singular *agyʎɔ* 'needle' which can also appear in the plural with that meaning).

Mass nouns tend to refer to matters (soil, flour, bread ...), or collective readings such as *la frytsɔ* 'fruits'. They generally resist pluralization, and show agreement in the singular, as shown in example 10.

(10) *la   'frytsɔ es   pu'liðɔ   u.'an*
     the.F.SG fruit   be.PRS.3SG beautiful.F.SG this.year
     'Fruits are of good quality this year'

When a plural form is forced, the interpretation is always that of kinds, except in those cases where the plural is lexical (e.g. *las farinetɔs* 'special dessert', from *farinɔ* 'flour').

There is only one noun with singular morphology but intrinsic plural value, *lu munde* 'people'. This word triggers singular agreement at the NP level (singular article and singular form of attributive adjectives, as shown in example 11), but plural agreement at the clausal level (example 12):

(11) *lu   pi'tju   'munde*
     the.MSG small.MSG people(M)
     'lower class people, the poor people'

(12) *lu   'munde   su   βeŋ'gyðes*
     the.MSG people(M) be.PRS.3PL come.PP.MPL
     'People have come'

There is historical evidence from closely related varieties that more nouns used to follow this pattern. The 16th century writer Auger Gaillard thus has plural agreement on the verb for singular subject nouns *canaillasso* 'bandits' or *rasso* 'race, group' (Nègre 1970:577). All of them referred to groups of persons with a singular noun triggering plural agreement at the clausal level, but singular agreement at the NP level.

Nominalized adjectives generally take the form of post-nominal adjectives, not that of prenominal adjectives, for those adjectives where there are two different forms: *lus pitjus*, not **lus pitjunes* 'the small ones, the kids'.

2.4 Verbal number

There is no evidence for verbal number in the language.

3 Agreement and the syntax of number

3.1 Agreement on verb forms

Finite forms of the verbs always agree in number with their subject. Different tense and mood combination, and different inflectional classes, have different ways of marking the distinction in number, but generally the marking of number is fused with the marking of person, and there is no specific marker of number: rather, each number and person combination makes use of a specific marker. There can be, in addition, some stem modification which largely depends on the tense, mood, person and number specification: root allomorphy is generally morphomic (see Maiden 2005 on morphomic distributions, and Esher 2012 for some cases in Occitan). Verbal morphology varies a lot between varieties, and it is not possible here for reasons of space to provide a complete survey. In Venés, for the present tense of regular conjugations, the forms are as shown in Table 4, not taking root allomorphy into account (when there are two forms in a cell, the first form is for the first conjugation, the second form in each cell for other conjugations). First and second person plural are characterized by having their stress on the ending.

Tab. 4: Person and number marking in the present tense in Venés.

	SG	PL
1	-i	-'an / -'ɛn
2	-ɔs / -es	-'as / -'ɛs
3	-ɔ / -Ø	-u

There are also periphrastic forms of the verb for perfect, pluperfect, past anterior, past subjunctive, pluperfect subjunctive, and past conditional which combine a finite form of the auxiliary with a past participle. In such cases, different verbs take a different auxiliary, either *aver* [a'βe] 'have' or *èstre* ['ɛstre] 'be'. The pattern of

agreement is different for each auxiliary. With auxiliary *èstre*, the auxiliary agrees in number with the subject, and the past participle agrees in number and gender with the subject of the clause, as shown in examples 13, 14 and 15:

(13) *'syj   βeŋ'gy-t*
    be.PRS.1SG come.PP-MSG
    'I(M) have come'

(14) *'su   βeŋ'gy-ðes*
    be.PRS.3PL come.PP-MPL
    'They(M) have come'

(15) *'su   βeŋ'gy-ðɔs*
    be.PRS.3PL come.PP-FPL
    'They(F) have come'

With auxiliary *aver*, the finite form of the auxiliary still agrees in number with its subject, but the past participle agrees in number and gender with the object of the verb:[5]

(16) *a'βɛm   man'tsaðɔj la-s   kas'taŋɔ-s*
    have.PRS.1PL eat.PP.FPL the.F-PL chestnut(F)-PL
    'We ate the chestnuts'

(17) *a'βɛm   man'tsaðɔ la   'supɔ*
    have.PRS.1PL eat.PP.FSG the.FSG soup(F)
    'We ate the soup'

In Venés, the paradigm of past participles is similar to that of prenominal adjectives (see 3.2.1), with the voicing of the stem final *-t* when it is followed by a vowel:

Tab. 5: Past participle inflection in Venés.

	SG	PL
M	-Ø	-es
F	-ɔ	-ɔs

---

5 In some varieties there are complex conditions licensing this agreement, mostly depending on the syntactic position of the object. In Venés, it seems that agreement occurs irrespective of the position of the object with regard to the verb, but may depend on the definiteness of the object.

3.2 Number agreement inside the NP

3.2.1 Morphology of number agreeing adjectives and other attributes

Adjectives agree in gender and number with their head noun inside the NP in most dialects. In Venés, there is an additional distinction in form in the masculine plural depending on whether the adjective is prenominal or postnominal (for those adjectives that allow both positions). Thus we find for the adjective *bɛl* 'beautiful, big' (M.SG) the forms '*bɛlɔ* (F.SG), '*bɛlɔs* (F.PL), but '*bɛles* (M.PL prenominal) and *bɛls* (M.PL postnominal). This also occurs for adjectives *pu'lit* 'pretty', *gran* 'big', *bu* 'good' and *pi'tju* 'small' with a similar inflectional paradigm, that is, only for adjectives that have final stress. It is thus possible that we have to consider that these adjectives have an overdifferentiated paradigm, with an extra cell compared to other adjectives. It is striking that prenominal adjectives in this variety have a similar paradigm to past participles, although past participles can never occur prenominally.

In Provençal, the distinction between prenominal and postnominal adjectives is that prenominal adjectives agree for both gender and number, while postnominal adjectives only agree in gender (see below 3.3). Thus an adjective such as *puli* 'pretty' (M.SG) has the forms *puli(z)* (M.PL, the last segment only surfacing prevocalically), *pulidɔ* (F.SG and F.PL when postnominal), and *pulidɛj* (F.PL prenominally). The overdifferenciation in Provence is thus on the feminine plural cell rather than the masculine as in Venés. This fact probably indicates a different origin for the two phenomena. Sibille (2016) indicates that in most Occitan varieties where the phenomenon is observed, one can trace its origins to the combined effects of final -s lenition and phrasal pitch accent falling on the last stressed syllable of the group. This type of explanation works well for the Provençal case, but fails to explain the two morphologically different forms in Venés, where final -s is very stable.

A number of varieties in the Pyrenées and the Alps have preserved a different type of marking for adjectives and modifiers, where the masculine plural is marked by *-i* instead of the expected *-s*. This has been shown to be a direct reflex of Latin second declension masculine plural subject marking, while other varieties have generalized the non-subject case form ending in *-s* (Sibille 2009).

3.2.2 The syntax of agreement

The question of number agreement inside the NP is complex. Different varieties differ as to what elements will agree with the noun in number. In some varieties, mostly Gascon and Languedocian varieties, as well as most Limousin and Auvergnat varieties, all members of the NP will agree in number and gender. In other varieties, such as Provençal, only prenominal elements agree in gender and num-

ber, postnominal ones only agreeing in gender. In a single variety do we find that only the noun is marked for plural and all other elements of the NP remain unmarked for number (but not for gender), in Alès. But beyond the obligatoriness of the number marking, a large number of varieties exhibit a differential treatment of prenominal and postnominal adjectives (see above 3.3.1).

In Venés, most members of the NP agree in number and gender with the head noun, and targets include all determiners (articles, demonstratives, quantifiers) and all prenominal and postnominal adjectives (with a slightly different paradigm each, see above 3.3.1). Numerals one and two agree in gender, and all numerals except *yn* 'one' trigger plural agreement. Prenominal and postnominal adjectives both agree, as in examples 18 and 19, but the numeral does not agree. Quantifiers agree with the head noun in gender and number, as in example 20, and some interrogative adjectives also agree as in example 21. Beyond the NP, there is number agreement with the verb for subjects, and with objects for periphrastic past tenses formed with the auxiliary *a'ße* 'have'.

(18) *la-s   'pawr-ɔj   'fennɔ-s*
      the.F-PL poor-FPL woman(F)-PL
      'The poor women'

(19) *la-s   'katre 'fennɔ-s   'pawr-ɔs*
      the.F-PL   four   woman(F)-PL poor-FPL
      'The four poor women'

(20) *i   a   'trɔp-ɔs   'trufɔ-s   per nu'zawtres*
      there have.PRS.3SG too.many-FPL potato-PL for us
      'There are too many potatoes for us'

(21) *'kynt-ɔj   βu'teʎɔ-j   'βɔj   dyr'bi?*
      which-FPL bottle(F)-PL want.PRS.2SG open.INF
      'Which bottles do you want to open?'

There are two main classes of adjectives, some that can only be postnominal, and some which can appear either before or after the noun, sometimes with a slight difference in meaning. For those adjectives that can be in either position, their paradigm is quite similar, except that they mark masculine plural in a different way: the prenominal form takes *-es*, the postnominal form only *-s* (see above 3.3.1). Here it is not possible to argue for a phonological explanation for the differential treatment of the form of the masculine adjective pre- and postnominally, as it can be for the Provençal resurfacing of a segment *-z* prevocalically for prenominal elements (see below).

There is no change in word order or in number with numeral modification. Numeral modification always triggers plural marking on the noun if the numeral used

is two or above. In that case, agreement inside the NP is similar to agreement when there is no numeral, and all determiners, some quantifiers and all adjectives agree in number with the modified noun. Similarly, if that modified NP is in a subject position, the verb agrees in number with it (example 23). When numerals are coordinated, it does not change the pattern of agreement. Numerals themselves are not marked for plural.

(22)   *'faj*   *me*   *pa'sa*   *'trez* *u* *'katre* *'tryfɔ-s*
      make.IMP.2SG 1SG.DAT pass.INF three or four   potato-PL
      'Give me three or four potatoes'

(23)   *j*   *a'βjɔ*   *dus* *'tipes* *k* *espe'raβu*   *da'βan*   *la*
      there have.IPRF.3SG two guys that wait.IPRF.3PL in.front.of the.F.SG
      'glɛjzɔ
      church
      'There were two guys waiting in front of the church'

Bare numerals can also act as predicates:

(24)   *'ɛru*   *'tres* *de* *fa'miʎɔ*
      be.IPRF.3PL three of family
      'They were a family of three (children)'

Non-natural numbers trigger different agreement depending on the type of numeral. Zero generally triggers plural agreement when elicited (speakers would naturally use a negative quantifier instead of zero in spontaneous data), as shown in example 25. On the contrary, 1.5 triggers singular agreement because of the syntactic construction used in that case (one cannot say 'one and a half potato' but have to say something like 'one potato and a half', example 26).

(25)   *'ɛj*   *ze'rɔ* *'tryfɔ-s*
      have.PRS.1SG zero potato-PL
      'I have zero potato'

(26)   *'ɛj*   *yn-ɔ* *'tryfɔ*   *e*   *'mjɛtsɔ*
      have.PRS.1SG one-F potato(F) and half
      'I have one and a half potato'

There is no occurrence in any of the surveyed dialects of Occitan of a plural word being used specifically to express the plurality of the noun they accompany, except if one considers articles to be such words in those dialects where nouns are un-

marked for number (Provençal). But articles are not there to specifically signal plurality, they also agree in gender and indicate the definiteness of the noun.

3.3 NP coordination

Two coordinated singular NPs in subject position normally trigger plural agreement on the verb. It is not possible to coordinate two NPs (with whatever number value) with a single determiner: each NP has its own determiner, which solves the question of potential number disagreements. When the coordinated NPs are of differing gender, default masculine gender is used on the past participle agreement. This is shown in examples 27 and 28:

(27) *a'kel   'ɔme e   a'kelɔ 'fennɔ su   mari'ðaðes*
     this.M.SG man and this.FSG woman be.PRS.3PL married.M.PL
     'This man and this woman are married'

(28) **a'kel / *a'keles   'ɔme e   'fennɔ*
     this.M.SG / this.M.PL man and woman

Similarly to articles, I couldn't find cases where an attributive adjective agrees with two conjunct nouns inside an NP. In cases where I tried to insist on it with my informants, they simply repeated the attributive adjective with each noun, as shown in examples 29 and 30:

(29) *l   'ɔme pi'tju   e   la   'fennɔ pi'tjunɔ*
     the.M.SG man small.M.SG and the.F.SG woman small.F.SG
     'The small man and the small woman'

(30) **l   'ɔme e   la   'fennɔ pi'tjus*
     the.M.SG man and the.F.SG woman small.M.PL
     'The small man and woman'

Thus there seems to be restrictions on the structure of coordinated NPs which prevent having a common modifier for two nouns. This restriction does not apply outside the NP.

4 Semantics and discourse

Regarding the pragmatic functions of number, the second person plural is used for politeness. Thus in addressing someone more senior, unknown, or of higher social

status, verbs of address will bear second person plural marking. But contrary to second person plural marking, the politeness form will not trigger plural agreement on past participles and predicative adjectives if addressing a single person in the polite form: these will remain in the singular, and agree in gender, and only agree in the plural when addressing more than one person.

(31)   *e*   *'βus,*   *i*   *'sjɛz*   *a'nat*   *al*   *mer'kat?*
      and you.(polite) there be.2PL gone.MSG to.the.M.SG market
      'And you(SG), did you go to the market?'

There is thus a clash in values in agreement for the polite use of second person plural agreement, which forms a distinct pattern than the normal use of second person plural. Regarding pronouns, if object and indirect object pronouns are the usual second person plural pronouns, the emphatic pronoun will have a special form, *bus*, distinct from the normal second person emphatic pronoun *buzawtres*.

    Because the language is pro-drop, most generic sentences will not use a pronoun. But second person reference will be used in most cases, just as is the case in the English "you never know". In most cases the form used will be second person singular, but it can happen that when the person who is talked to is more senior or of higher social status the second person plural will be used instead, alongside a possible second person singular. There is another case of genericity, for verbs of saying and thinking, which refers to what is commonly thought or said by people: in such cases, an indeterminate third person plural will be used on the verbs, as in example 32:

(32)   *'dizu*   *ke* *lu*   *'mɛrɔ* *se*   *turna'ra*   *pas*
      say.PRS.3PL that the.M.SG mayor REFLX do.again.FUT.3SG not
      prezen'ta
      present.INF
      'People say that the mayor will not run in the next election'

Generic sentences (*beavers build dams* – cf. *each beaver builds dams*) and reference to kinds (*bears are extinct* – cf. **Each bear is extinct*) both use the definite plural. It is not possible in generic readings to use a determinerless NP or a partitive, as shown in example 34 where the use of the partitive cancels a generic reading present with the definite determiner to make reference to specific cows:

(33)   *laj*   *'βakɔ-s* *'mantsu*   *d*   *'ɛrbɔ*
      the.FPL cow-PL eat.PRS.3PL PRTV grass
      'Cows eat grass' (generic reading)

(34)   *de* *'βakɔ-s* *'mantsu*   *d*   *'ɛrbɔ*
      of cow-PL eat.PRS.3PL PRTV grass
      'Some cows eat grass' (specific reading)

If count nouns appear in the plural in generic sentences, mass nouns will on the contrary appear in the definite singular. A plural form is impossible because it would produce a clash between the specific, sorts reading of the plural of mass nouns, and the genericity given by the use of the definite article.

(35) *lu*   *βi*   *'blank*   *'kal*   *'pas que su.'εsɔ*
     the.M.SG wine white.M.SG have.to.PRS.3SG not that be.IPRF.SBJV.3SG
     'trɔk   *'kaut*
     too.much hot.M.SG
     'White wine should not be drunk too hot'

In the usual case, all plural nouns will be count nouns. The only exception is that mass nouns can be pluralized and receive a sorts reading. This overall seems to be extremely rare in spontaneous speech. Mass nouns can only allow quantification and numeral modification in their mass singular form accompanied by a classifier which receives the plural interpretation, as shown in example 36 (*blat* is a mass noun).

(36) *seme'naβεn*   *'trej*   *βarie'tats de 'blat*   *a l*   *e'pɔkɔ*
     sow.IPRF.1PL three varieties of wheat at the.F.SG time
     'We used to sow three varieties of wheat at the time'

When used in the plural, mass nouns can no longer take mass determiners (no possibility of use with the partitive), which shows that in such cases they are no longer considered as mass nouns.

     Questions and negative context allow for singular reference of a plural, much in the same way it does in English. Interestingly, there seems to be an effect with animacy here, in that inanimates allow the reading 'one' as an answer to a plural question (example 37), but it does not seem to be possible for animates, potentially because *lus pitjus* is here understood as a coherent group (example 38). Venés Occitan does not allow bare plurals in object position (this varies a lot between dialects, Gascon allowing bare plurals, Languedocian presenting an articleless partitive, Limousin a partitive with the article, like French does).

(39) - *Esk a'βεn*   *de*   *'pumɔ-s?*
      Q have.PRS.1PL PRTV apple-PL

    - *n*   *a'βεm*   *pas k*   *'ynɔ*
      of.it have.PRS.1PL not that one.F
     '- Do we have apples? - We only have one'

(40) - *Su   arri'ßað-ej   lu-s   pi'tju-s?*
     be.PRS.3PL arrived-MPL the.M-PL small.M-PL
    - *'Nani j   a   pas k   A'lex k   ez   arri'ßað-ɔ*
     No   there have.PRS.3SG not that Alex that be.PRS.3SG arrived-FSG
    '- Have the little ones arrived? - No, only Alex arrived'

Under negation, there can sometimes be a similar effect (in some cases, the negative quantifier requires a noun in the singular and thus does not trigger such effects):

(41) - *j   a   paj 'mait   de   'luts   aj'si*
     there have.PRS.3SG not anymore PRTV woolves here
    - *si   ke   ne   ße'tsɛr-u   'yn   a La'kɔwnɔ*
     yes that of.it see.PST-3PL one.M.SG in Lacaune
    '- There are no wolves anymore here. - Yes there are, they saw one in Lacaune'

5 Conclusions

The morphological realization of number in Occitan is highly variable across dialects. Some varieties mark number throughout the NP, other restrict it to prenominal elements, some mark it uniformly with a final -s or one of its allomorphs (Venés), others present inflectional classes (Usseaux), and yet others use non segmental realizations (Limousin). In a number of varieties there is an asymmetry between prenominal and postnominal adjectives, which present a different paradigm. In Provençal this asymmetry affects all prenominal elements, marked for number, while nouns and postnominal modifiers remain unmarked. The variability in the marking of number can be ascribed to the differential treatment of unstable -s in diachrony for different varieties.

Plural number seems to have a simple basic meaning of plurality (more than one), except in certain very specific discourse contexts such as questions and negation where the plural can sometimes receive a non-plural reading.

Pronominal number is marked through suppletive forms in the first and second person while nominal and adjectival number, at least in Venés, is marked segmentally.

The semantics and discourse functions of number have never been studied before in the language, and would largely benefit from further research. In particular, it would be interesting to see if varieties that do not mark number on nouns still have the same semantic readings.

Abbreviations

1	first person
2	second person
3	third person
DAT	dative
F	feminine gender
FUT	future
IMP	imperative
INF	infinitive
IPRF	Imperfect
M	masculine gender
NP	noun phrase
PL	plural
PP	past participle
PRS	present
PRTV	partitive
PST	past
Q	question marker
REFLX	reflexive
SBJV	subjunctive
SG	singular

References

Amaro-Péguy, Lucie. 2014. *L'occitan alpin d'Usseaux. Description d'une langue en danger, en contact avec deux aires dialectales (francoprovençale et piémontaise), et sous l'influence de deux langues standards (français et italien)*. PhD Thesis, Université Lumière Lyon 2.

Bach, Xavier. 2012. The Morphology of Number on the Occitan Noun Phrase. MSt Thesis, University of Oxford.

Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: SEDES.

Floricic, Franck. 2010. Remarques sur le marquage du nombre dans le parler occitan de Veyrines-de-Vergt. In Franck Floricic (ed.), *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*. Paris: ENS Éditions.

Jagueneau, Liliane. 1979. *Recherche sur l'opposition singulier-pluriel en occitan central*, Thèse de 3e cycle, Poitiers: Université de Poitiers.

Javanaud, Pierre. 1981. *The vowel system of Lemosin. A phonological study*. Göteborg: University of Göteborg.

Lieutard, Hervé. 2004. Spécificité morphologique du pluriel languedocien : la notion de "cheville". *Cahiers de Grammaire* 29. 89–104.

MacKenzie, Laurel. 2010. /s/-Deletion and the Preservation of Plurality in Modern Occitan. *University of Pennsylvania Working Papers in Linguistics* 16–2 'Selected Papers from NWAV 38' [retrieved online http://repository.upenn.edu/pwpl/vol16/iss2/15].

Mok, Q. I. M. 2008. L'opposition singulier/pluriel dans le parler de quelques communes du Périgord Blanc. In François Pic & Patrick Sauzet (eds.), *Per Q. I. M. Mok. Études de linguistique occitane moderne*, 3–16. Aachen: Shaker.

Nègre, Ernest. 1970. *Auger Gaillard. Oeuvres complètes*. Paris: Presses Universitaires de France.

Quint, Nicolas. 1996. *Grammaire du parler occitan nord-Limousin marchois de Gartempe et de Saint-Sylvain-Montaigut (Creuse). Etude phonétique, morphologique et lexicale.* Limoges: La Clau Lemosina.

Quint, Nicolas. 1998. *Le parler occitan alpin du Pays de Seyne : Alpes de Haute-Provence.* L'Harmattan.

Ronjat, Jules. 1937. *Grammaire istorique des parlers provençaux modernes. Tome III.* Montpellier: Société des Langues Romanes.

Sauzet, Patric. 2011. Los morfèmas de plural nominal a Sant Júlia de Cremsa : [-w] e lo ton bas. In Angelika Rieger (ed.), *Actes du 9e Congrès de l'Association Internationale d'Etudes Occitanes,* 827–842. Aachen: Shaker Verlag.

Sauzet, Patrick. 2012. Occitan Plurals: A Case for a Morpheme Based Morphology. In Sascha Gaglia & Marc-Olivier Hinzelin (eds.), *Inflection and word formation in Romance languages,* 179–200. Amsterdam: Benjamins.

Sauzet, Patrick & Michèle Oliviéri. 2016. Southern Gallo-Romance (Occitan). In Martin Maiden & Adam Ledgeway (eds.), *The Oxford Guide to the Romance Languages*, 319–349. Oxford: OUP.

Sibille, Jean. 2009. Les formes en -i issues du nominatif pluriel de la 2ème déclinaison latine, en occitan : essai d'approche panchronique. In Claudine Fréchet (ed.), *Langues et cultures de France et d'ailleurs*, 233–250. Lyon: Presses Universitaires de Lyon.

Sibille, Jean. 2011. La marca del numero nella parlata occitana di Sénaillac-Lauzès (Francia). *Rivista Italiana di Dialettologia* 35. 165–184.

Sibille, Jean. 2015. *Description de l'occitan parlé à Sénaillac-Lauzès (Lot) et dans les communes voisines*. Limoges: Lambert Lucas.

Sibille, Jean. 2016. Morphologie nominale et morphologie pré-nominale dans plusieurs variétés d'occitan. In Franz Rainer, Michela Russo & Fernando Sánchezz Miret (eds.), *Actes du XXVIIe Congrès international de linguistique et de philologie romanes (Nancy, 15–20 juillet 2013). Section 3 : Phonétique, phonologie, morphophonologie et morphologie*, 163–174. Nancy: ATILF.

Adrian Stegovec

7 Number in Slovenian

Abstract: Slovenian is one of the very few modern Indo-European languages with a productive three-way number opposition across all major categories. Additionally, the expression of number is part of a complex fusional inflectional system; on nouns alone, the three-way number opposition interacts with the expression of three grammatical genders, six cases, and inflection class differences. The morphological patterns that emerge display a large amount of number syncretism, number-conditioned syncretism, as well as interesting patterns of number-conditioned suppletion and allomorphy. In the realm of syntax, quantified noun phrases display an extremely complex pattern of number and case marking interactions, which in turn have consequences for verbal agreement, including mismatches in number agreement between auxiliary verbs and participles and complex patterns of agreement with conjoined noun phrases. Finally, the three-way number system has interesting consequences for the lexical distribution and semantics of number on nouns, in particular cases of seemingly optional dual number.

1 Overview

Slovenian, also called Slovene,[1] is a Slavic language from the South Slavic branch. It has roughly 2.4 million speakers, most of whom live in the Republic of Slovenia and border areas in Austria, Italy, and Hungary. Perhaps the most striking property of Slovenian in the domain of grammatical number is its productive three-way number system distinguishing singular, plural, and dual number. Slovenian, Upper and Lower Sorbian are the last Slavic languages to retain the productive dual from Proto-Slavic (Schenker 1993: 85).[2] Outside these three closely related languages, dual persists only vestigially in a handful of modern Indo-European languages.[3]

---

1 There is some debate about whether 'Slovenian' or 'Slovene' should be the default English name for the language, although both are perfectly acceptable (Klinar 1996; Mileusnić 2009). I use 'Slovenian' exclusively in this chapter, mainly due to it being more common by a significant margin (Mileusnić 2009: 55–6) and due to it gradually becoming the predominant name used by international and Slovenian linguists.
2 Kashubian (also Cassubian) was probably the last Slavic language to lose the dual: Slovincian (extinct by some time after the 1950s) and dialects immediately to its east still had dual at the beginning of the 20th century, while some northern dialects of Kashubian retained dual in first person pronouns until the 1950s (Stone 1993: 768). More recent works on Kashubian, like Hopkins (2001), list only singular and plural forms (see also Slobodchikoff 2013, 2019).
3 Of the few modern Indo-European languages identified as having a dual by Humboldt ([1828] 1997), Lithuanian only has dual forms for 'two' and 'both' and some dialects also have dual forms for certain pronouns and nouns, but only if modified by 'two' or 'both' (Ambrazas 2006: 102, 184–

It is therefore not surprising that Slovenian has received much linguistic attention in relation to its dual, starting with Lucien Tesnière's seminal *Atlas linguistique pour servir à l'étude du duel en slovéne* [Linguistic atlas to be used for the study of the dual in Slovenian] (Tesnière 1925a) and the accompanying monograph *Les formes du duel en slovène* [Forms of dual in Slovenian] (Tesnière 1925b), which remain among the most influential works on the Slovenian language. Some more recent works on the Slovenian dual which provide a comprehensive overview of the subject (in English) include Derganc (1994, 2003, 2006), Jakop (2008, 2012, 2020) and Marušič and Žaucer (2021), see references for updated entry. The three-way number system will inevitably figure into any discussion of number in Slovenian, but it is not the goal of this chapter to just be another overview of the Slovenian dual. Slovenian has many grammatical properties – many of them characteristic of Slavic languages – that interact with number, either directly or indirectly, such as: rich inflectional morphology, a rich case system, a relatively complex agreement system, and a relatively free word order of major syntactic constituents. Let us briefly review how number plays a part in these phenomena and why examining the role of number in each of them is relevant for linguistic theory.

Number is expressed in Slovenian through fusional inflectional suffixes which encode simultaneously the value of several inflectional categories; like most Indo-European languages, Slovenian does not have a grammaticalized classifier system. Aside from number, the inflectional suffixes may also express person, gender, animacy, and case information. Which and how many categories can be simultaneously expressed by the suffix depends on the type of stem and the value of the individual categories. The three-way number system plays a prominent role in these interactions: number is both a target and trigger of *syncretism* as well as a trigger for *stem suppletion*. We will see that the relevant Slovenian data are challenging for existing analyses of both phenomena.

Number also plays a key role in syntactic phenomena like case and agreement. The type of quantifier or numeral modifying a noun phrase – as well as its number – determines its case marking. The case of the quantified noun phrase then in turn also affects agreement, both internally to the noun phrase and with the noun phrase. The relatively free word order in Slovenian also plays a role in constraining (number) agreement with conjunctions: the position of the conjunction in relation to an agreeing element determines whether the first, last, or both conjuncts can be agreed with. Together, these options result in a highly complex system of syntactic dependencies.

5, 215–7), Scottish Gaelic has dual forms for some short feminine nouns only if they are modified by 'two' (Lamb 2008: 206), and Welsh likewise has dual forms for very few nouns only if modified by 'two' (King 2016: 73–4). Outside Humboldt's list, some dialects of Northern Frisian have been reported as having dual pronouns, although they are disappearing from everyday speech (Howe 1996: 193–5), while Breton has recently developed a dual (MacAulay 1992: 416–7), although only in "pair nouns" (see Section 2.3.3).

Finally, the three-way number system is interesting from the perspective of the lexical distribution and semantics of individual number values. For example, there are *pluralia tantum* but no *dualia tantum* nouns in Slovenian, and while bare plurals allow a generic reading, bare duals seem not to be able to. The issues concerning lexical number and the interpretation of number intersect with a class of nouns that can appear plural, but nonetheless denote only a pair of entities.

The chapter is organized so that number in pronouns and nouns are considered first, with an emphasis on the morphological expression and lexical distribution of number (Section 2). The role of number in syntactic dependencies is then examined by looking at adjectives (and other elements that can modify noun phrases) and verbs, where the effects of number on case and agreement can be seen (Section 3). Lastly, the semantic and pragmatic contribution of the three number values is considered, highlighting the differences and similarities between them (Section 4).

A note on the Slovenian socio-linguistic situation is on order before we start. Slovenian shows a great deal of dialectal variation, with more than 40 regional varieties traditionally divided into seven major dialect groups: (i) Carinthia (*Koroška*), (ii) Littoral (*Primorska*), (iii) Rovte, (iv) Upper Carniola (*Gorenjska*), (v) Lower Carniola (*Dolenjska*), (vi) Styria (*Štajerska*), and (vii) Pannonian (*Panonska*) (Logar and Rigler 2016; Herrity 2016: 2–3), and with significant grammatical differences between many of them. The official standard used in education, mass media, and on formal occasions is not based on any currently spoken regional dialect, which makes all speakers functionally bidialectal (Tollefson 1978: 94–100; Herrity 2000: 1–4, 2016: 1–14). This creates a lot of free grammatical variation in the colloquial standard variety, which differs both from the official standard and specific regional varieties, and serves as a kind of lingua franca in informal situations. A detailed examination of the dialectal differences related to number is impossible in this chapter (but see Tesnière 1925a,b; Jakop 2008, 2012, 2020, which focus on the dual), however some of the differences and examples of free variation are very instructive and will be noted when relevant. Unless noted otherwise, the provided examples are from the colloquial standard variety.

2 Pronominal, nominal, and verbal number

2.1 Generalities

In Slovenian, only pronouns (Section 2.2) and nouns (Section 2.3) have inherent number values, although some types of pronouns and specific nouns lack number inflection, some nouns are number deficient, and several morphological processes can neutralize specific number oppositions. Slovenian has no grammaticalized system of verbal number (Section 2.4).

2.2 Pronominal number

Slovenian pronouns fall into three classes in terms of number marking: (i) pronouns with full number marking (this section and Section 2.2.2); (ii) pronouns whose inflectional paradigm is limited based on number (Section 2.2.3); and (iii) number-invariant pronouns (Section 2.2.4).

Personal pronouns differ from nouns and adjectives in that the former maintain the three-way number contrast through the whole case paradigm (see Section 2.2.2 for details). The inflection pattern of strong (tonic) personal pronouns (clitic/weak pronouns are discussed in Section 2.2.3) is illustrated in Table 1 (inflectional morphology is in bold font; the order in which number values are given in all tables is: singular, plural, dual – the reason for this will become clearer below).[4]

Tab. 1: Personal pronouns in Slovenian.

		NOM	ACC	GEN	LOC	DAT	INS
SG	1	jaz	men**e**		men**i**		man**o**[5]
	2	ti	teb**e**		teb**i**		tab**o**
	3.M	on	nje**ga**		nje**mu**[6]		nj**im**
	3.N	on**o**					
	3.F	on**a**	nj**o**	nj**e**	nj**ej**		nj**o**
PL	1	m**i**	na**s**			na**m**	na**mi**
	2	v**i**	va**s**			va**m**	va**mi**
	3.M	on**i**	nj**ih**			nj**im**	nj**imi**
	3.N	on**a**					
	3.F	on**e**					

---

4 Grammatical information in tables and glosses is abbreviated following the Leipzig glossing rules, with the addition of: AUG 'augment', CNTR 'contrastive (topic) particle', DER 'derivational', and PTCL 'particle'. The order of cases in the tables follows Caha (2009) (see also Tesnière 1925b) instead of the NOM-GEN-DAT-ACC-LOC-INS order adopted in recent Slovenian linguistics (Toporišič 2000).
5 1st and 2nd person singular have two instrumental variants: *mano~menoj* and *tabo~teboj* respectively. The forms in Table 1 are the colloquial variants; the alternative forms are very literary (Herrity 2000: 90, 2016: 125).
6 An alternative 3rd person M/N locative form is *njem* (more parallel to the adjectival inflection; see Table 6); this is also the prescriptively enforced form (Toporišič et al. 2001). For ease of exposition, I only present the locative forms shown in Table 1 (which are the ones used in my variety of colloquial Slovenian).

Tab. 1: Continued.

		NOM	ACC	GEN	LOC	DAT	INS
DU	1.M	mid**va**	na**ju**			na**ma**	
	1.F	mid**ve**					
	2.M	vid**va**	va**ju**			va**ma**	
	2.F	vid**ve**					
	3.M	ona**dva**	nji**ju**[7]			nj**ima**	
	3.F/N	one**dve**					

Notice that the nominative dual forms are constructed by adding the numeral *dva* 'two'. For example, the 1st person dual form can be decomposed into the pronominal base *mi* and the numeral *dva*. The numeral is obligatory in nominative forms, whereas with oblique forms it is only optionally added for emphasis (e.g. *na**ju** dv**a*** 'us.DU.ACC two.DU.M.ACC'; Toporišič 2000: 306). Both parts of such dual pronouns are inflected: the numeral part is inflected like an adjective (see Section 3.1 for details), while the pronominal base is inflected like a regular personal pronoun.

The numeral part of nominative dual pronouns consistently inflects for gender, distinguishing masculine pronouns (-*dv**a***) from feminine and neuter ones (-*dv**e***).[8] This is not always the case with the pronominal part (see Table 1): 3rd person forms make the same gender distinctions as the numeral part (*ona*- '3.DU.M-' vs. *one*- '3.DU.F/N-'), whereas 1st and 2nd person forms, which are also identical to the plural forms, do not. There is also some variation in gender marking with dual and plural pronouns: the nominative plural and dual forms shown in Tables 2 and 3 respectively exist as variants alongside the forms shown in Table 1; the two versions are in free variation in standard Slovenian (Toporišič 2000: 305–6, Herrity 2000: 89–90, 2016: 124–5).[9]

---

7 The *nji**ju*** form can be simplified to *n**ju*** in colloquial speech (sometimes even in written form, as a quick online search reveals). The simplified form also appears in preposition pronouns (see Section 2.2.3).

8 Notice that 1st and 2nd person pronouns never have neuter forms. Since such pronouns must express the semantic (natural) gender of the speaker or addressee respectively, their gender can only be masculine or feminine.

9 There are dialectal differences in the use of the variants; e.g. in my regional dialect (Nova Gorica; Littoral dialect group) the Table 1 variants are strongly preferred. See Jakop (2008: Ch. 5) regarding dialectal variation in dual pronouns.

Tab. 2: Alternate NOM plurals.

		NOM
PL	1.M	m**i**
	1.F	m**e**
	2.M	v**i**
	2.F	v**e**
	3.M	on**i**
	3.N	on**a**
	3.F	on**e**

Tab. 3: Alternate NOM duals.

		NOM
DU	1.M	m**i**dv**a**
	1.F	m**e**dv**e**
	2.M	v**i**dv**a**
	2.F	v**e**dv**e**
	3.M	on**adv**a
	3.F/N	on**i**dv**e**

With the Table 1 variants, the segments shared by plural and dual pronouns are gender-invariant in 1st and 2nd person (*mi*(-) '1NSG.M/F', *vi*(-) '2NSG.M/F'). In contrast, the corresponding forms in Tables 2 and 3 show gender distinctions with all three persons. I will return to discuss some implications of these patterns of number-gender interaction in Section 2.2.2.

Another point of variation is with dual locative forms: instead of the forms in Table 1, which are identical to their accusative and genitive counterparts (*naju/ vaju/njiju*), some speakers use locative forms identical to dative and instrumental forms (*nama/vama/njima*) (see also Section 2.2.2).

Apart from these minor morphological differences, dual pronouns are very stable across dialects. As noted already by Tesnière (1925a,b) and confirmed by Jakop's (2008) study, pronouns are resistant to the loss of dual morphology even in the dialects where gradual loss of the dual is taking place, as those retain at least dual nominative pronouns. The loss of oblique dual forms is limited to south-west Littoral dialects and south-eastern Lower Carniola dialects (Jakop 2008: Ch. 5, 2012: 355). In these dialects, the addition of the 'two' to oblique nonsingular pronouns (e.g. *nas dva* 'us.NSG.ACC two.DU.ACC') is often used to disambiguate the dual and plural uses of the pronouns.

2.2.1 Suppletion in personal pronouns

The status of stem suppletion in personal pronouns is controversial (see Corbett 2005 for an overview of the main issues). Pronoun stems that assume different forms depending on the values of specific inflectional categories, such as number, can be analyzed as suppletion (by analogy with such cases in nouns), as distinct lexemes in complementary distribution, or as either of the two depending on the specific example (see Mel'čuk 1994: 387–8 regarding the last option).[10] I will assume here

---

10 I adopt Aronoff's (1994) terminology in the morphological discussions: a *lexeme* is an abstract morphological unit linking form, meaning, and syntactic function that is unspecified for the contextually variable categories encoded by inflection, a *stem* is the realization of a lexeme to which a

Tab. 4: Personal pronoun stems.

		NOM	ACC	GEN	LOC	DAT	INS
SG	1	jaz	m-<u>en</u>				m-<u>an</u>
	2	ti	t-<u>eb</u>				t-<u>ab</u>
PL/DU	1	m	na				
	2	v	va				
SG/PL/DU	3	on	<u>nj</u>				

that such irregular form alternations in personal pronouns involve stem suppletion; see Corbett (2005), Bobaljik (2015), Moskal (2015a,b), Smith et al. (2019), among others, for arguments that including personal pronouns in the typology of suppletion can lead to important theoretical insights.

In Slovenian, we can isolate the personal pronoun stem by excluding the inflectional suffixes (and numerals), which are mostly identical to those of nouns and adjectives (see Section 2.2.2). The pronoun stems, shown in Table 4, mainly encode person contrasts; the underlined segments are "support morphemes" that differentiate strong personal pronouns from their clitic counterparts (see Section 2.2.3). Two possible analyses of the suppletion pattern are suggested by the dashed lines. The more conservative analysis treats the changes in the support morpheme and vowel alternations as suppletion, while the less conservative one treats those as regular morphological changes. Crucially, neither option affects the number suppletion pattern.

Notice that the stem suppletes for number: singular versus nonsingular with 1st and 2nd person, as well as case: nominative versus oblique cases (1/2SG.INS involves an additional change limited to the support morpheme).[11] The presence of number suppletion with 1st and 2nd but not 3rd person could be seen as a type of *Animacy Hierarchy* effect (Corbett 2000: Ch. 3–4): a given number opposition must affect a top segment of the scale in (1). In other words, the number opposition cannot exist in a lower segment of (1) if it does not also exist in all higher segments.

(1)　<u>1st person > 2nd person</u> > 3rd person > kin > human > animate > inanimate

---

given affix is attached or upon which a given nonaffixal realization rule operates, and a *root* is the default form associated with a lexeme that is morphologically unanalyzable and abstracts away from all morphological structure (see Aronoff 1994: 39–40). Although Aronoff's terminology is by no means theory neutral, I find that it is most easily translatable into related concepts in other frameworks.

11 With the alternative instrumental forms, *me<u>no</u>j* and *te<u>bo</u>j* (see footnote 5), this change does not occur. In fact, even with *ma<u>no</u>* and *ta<u>bo</u>*, the vowel mutation on the support morpheme could be analyzed as ablaut triggered by the suffix.

In the case of the number suppletion pattern in Table 4, there are more number distinctions expressed via stem suppletion with pronouns high on (1), namely 1st and 2nd person pronouns.

2.2.2 Syncretism in personal pronouns and beyond

Slovenian personal pronouns display a lot of *syncretism* in their inflection – there are much fewer inflectional suffix forms than underlying combinations of inflectional category values (see Baerman, Brown, and Corbett 2005 and Caha 2019 for overviews of the main theoretical issues concerning syncretism). Crucially, it is not only number distinctions that are syncretic in the context of other categories in Slovenian (cf. English number syncretism in 2nd person forms: *you* is both the singular and plural form), specific number values can also be the context for syncretism in other inflectional categories. Because the relevant syncretism patterns show significant overlap – as well as some key differences – across different syntactic categories, I will situate the discussion of syncretism in personal pronouns within the broader context of syncretism in other related inflection patterns.

Let us first establish that number-conditioned syncretism exists in personal pronouns. As seen in Table 5, there is more case and gender syncretism in nonsingular suffixes than in singular suffixes, but also more case syncretism in dual suffixes than in plural and singular suffixes.

Tab. 5: Personal pronoun inflection.

		NOM	ACC	GEN	LOC	DAT	INS
SG	3.M	-Ø	-ga		-mu		-im
	3.N	-o					
	3.F	-a	-o	-e	-i		-o
	1/2						
PL	3.M	-i	-ih			-*V*m	-*V*mi
	3.N	-a					
	3.F	-e					
	1/2		-s				
DU	3.M	-a -a	-ju			-*V*ma	
	3.N	-e -e					
	3.F						
	1/2	-i -a					

Tab. 6: Adjectival inflection.

		NOM	ACC	GEN	LOC	DAT	INS
SG	3.M	-Ø	-ga		-em	-emu	-im
	3.N	-o					
	3.F	-a		-e	-i		-o
	1/2						
PL	3.M	-i	-e	-ih		-im	-imi
	3.N	-a					
	3.F						
	1/2						
DU	3.M	-a				-ima	
	3.N	-i					
	3.F						
	1/2						

To see that number distinctions themselves may be syncretic in Slovenian, we need to look at nouns and adjectives. The comparison between the inflectional suffixes of personal pronouns, nouns, and adjectives is particularly instructive here in revealing the role of the three-way number system in the distribution of syncretic forms due to the significant overlap in the suffixes used across the three syntactic categories. The adjectival inflection pattern is illustrated in Table 6 and the nominal inflection pattern is illustrated in Table 7;[12] some suffix forms given in the tables abstract away from changes in certain vowels (represented as "*V*"), as the presence and quality of those vowels is governed by the inflection class of the stem and its phonological properties (see Toporišič 2000: 273–301, 321–5; Herrity 2000: 40–70, 72–3, 2016: 61–95, 99–101).

Notice that there is a consistent plural/dual syncretism across genitive and locative forms with both adjectives and nouns. This syncretism pattern can be characterized as in (2).

---

12 There is more variation not shown in Table 7, namely another instance of case and number syncretism in Class III nouns (see Section 2.3.1). Inflection class takes precedence over gender in determining the inflection pattern for nouns, as I discuss in Section 2.3.1. However, as there are no class effects in personal pronouns and adjectives, I compare the paradigms based on gender here. Crucially, everything I say about gender syncretism below can also be stated in terms of class (see footnote 15). Related to this, I only consider animate adjectival and nominal inflection here (the difference is confined to M.SG.ACC forms; see Sections 2.3.1 and 3.1), since personal pronouns always take animate inflection.

Tab. 7: Nominal inflection.

		NOM	ACC	GEN	LOC	DAT	INS
SG	3.M	-Ø	-a		-u		-Vm
	3.N	-o					
	3.F	-a		-e	-i		-o
	1/2						
PL	3.M	-i	-e	-Vv	-Vh	-Vm	-V(m)i
	3.N	-a		-Ø			
	3.F						
	1/2						
DU	3.M	-a				-Vma	
	3.N	-i					
	3.F						
	1/2						

(2) ***Case-conditioned number syncretism:***
The expression of *plural number* is identical to the expression of *dual number* in the environment of *nouns* and *adjectives* marked with *genitive* and *locative case*.

In contrast, number-conditioned case and gender syncretism is consistent across the three inflection patterns, with variation in the amount of gender syncretism. However, all three syntactic categories show dative and instrumental syncretism with dual number, which can be characterized as in (3).

(3) ***Number-conditioned case syncretism:***
The expression of *dative case* is identical to the expression of *instrumental case* in the environment of *dual number*.

Recall though from Section 2.2 that there is speaker variation regarding case syncretism in dual personal pronouns – specifically, the **-ima** suffix can span locative, dative, and instrumental case instead of only dative and instrumental case. This means that these speakers have an alternative case syncretism pattern in personal pronouns only, which can be characterized as in (4) (the equivalent suffix in nouns and adjectives still patterns according to (3)).

(4) ***Number-conditioned case syncretism – some speakers:***
The expressions of *locative*, *dative*, and *instrumental case* are all identical in the environment of *dual number* in *personal pronouns*.

The syntactic category of the stem is relevant also in the case of total gender syncretism. Personal pronouns show total gender syncretism with nonsingular numbers in all oblique cases, adjectives show it in four oblique cases (genitive, locative, dative, and instrumental), and nouns show it in only three oblique cases (locative, dative, and instrumental). The three patterns can be characterized as in (5); note that the number component of the conditioning environment is constant across all three.

(5) **Number and case-conditioned gender syncretism**
   The expressions of *masculine*, *feminine*, and *neuter gender* are identical in the environment of *plural* and *dual number* in ...
   a. *personal pronouns* marked with *oblique (non-nominative) cases*;
   b. *adjectives* marked with *genitive*, *locative*, *dative*, and *instrumental case*;
   c. *nouns* marked with *locative*, *dative*, and *instrumental case*.

To sum up, the roles number plays in syncretism are: (i) number can be syncretic in the environment of other inflectional categories (cf. (2)), (ii) number can be the environment for syncretism of other inflectional categories (cf. (3)), and (iii) number can together with another inflectional category be the environment for syncretism of a third inflectional category (cf. (5)).

The specific examples of (i) and (ii) we considered above show, respectively, that the expression of number distinctions depends on the case and that the expression of case distinctions depends on the number. As Corbett (2000: 275) points out, the existence of both (i) and (ii) in Slovenian is surprising in light of Aikhenvald and Dixon's (1998) typological study of syncretism, which did not find any languages where such dependencies hold in both directions (see Aikhenvald and Dixon 1998: 63n20). Aikhenvald and Dixon speculate that, if found, such a state of affairs would be transitory and the result of phonological changes. There are, however, several instances of number-conditioned case syncretism in Slovenian in all three relevant syntactic categories (as seen in Tables 5–7),[13] and if we examine closely the one example of case-conditioned number syncretism, we see that it is not tied to any specific suffix form: with adjectival inflection (see Table 6), the genitive/locative suffix **-ih** is the only plural/dual syncretic form, while with nominal inflection (see Table 7) we observe plural/dual syncretism in three separate suffixes: **-Vh** (locative), **-Vv** (genitive masculine), and **-Ø** (genitive feminine/neuter); and there is even more variation if we compare different inflection classes (see Section 2.3). The per-

---

13 Excluding speaker variation (cf. (4)), the examples of number-conditioned case syncretism are: DAT/INS syncretism in dual (personal pronouns, adjectives, nouns), NOM/ACC syncretism in dual (adjectives, nouns), ACC/GEN/LOC syncretism in dual and plural (personal pronouns), GEN/LOC syncretism in dual and plural (adjectives), and – if we count unmarked singular cases – LOC/DAT syncretism in singular (personal pronouns, nouns).

vasiveness of number-conditioned case syncretic forms and the fact that the case-conditioned number syncretism is tied to the underlying distinctions rather than a specific suffix is at least highly suggestive that the bi-directional case-number dependency in Slovenian is not just due to a surface phonological quirk (see also Baerman, Brown, and Corbett 2005: 163–6 for a relevant discussion that also touches on Slovenian).

I should note before moving forward that in addition to the many cases of number-conditioned case syncretism beyond the one described in (3)–(4), there are also cases of number-conditioned gender syncretism beyond the total syncretism described in (5). I will briefly discuss one of the other gender syncretisms later in this section in the context of Nevins' (2011) analysis of Slovenian. However, closely examining any other cases would take us too far afield.[14]

The descriptions of the syncretism patterns that were offered so far, while describing the facts, do not reveal anything about what unifies, for example, the environments in which plural and dual are syncretic, or the syntactic categories in which the syncretism is observed. Furthermore, when we take a closer look at the number and case-conditioned gender syncretism (cf. (5)), we observe a consistent pattern that can be summarized using the scale in (6): there is more gender syncretism with the cases high on the scale than with those low on the scale.

(6) instrumental/dative/locative > genitive > accusative > nominative

We would not expect such patterns to emerge from an arbitrary assortment of syntactic categories and cases. We would, however, expect such patterns to emerge if specific values of the relevant categories form natural classes. In their seminal works, Jakobson ([1936] 1984a, [1958] 1984b) and Bierwisch (1967) argued that these natural classes should be formally captured by decomposing categories like *plural*, *accusative*, and *neuter* into more primitive abstract oppositions – encoded by way of binary grammatical features (see Calabrese 1998, 2008 for an updated take on this view).

---

14 Two case syncretisms worth a brief mention, due to being quite prevalent, are (i-a) and (i-b); with the adjustment for personal pronouns in (ii-a) and (ii-b) with the LOC/DAT/INS syncretism alternative (cf. (4)) (see also footnote 16).

(i) ***Number-conditioned case syncretism***

    a. The expressions of *accusative*, *genitive*, and *locative case* are identical in the environment of *dual* and *plural number* in *personal pronouns*;

    b. The expression of *genitive case* is identical to the expression of *locative case* in the environment of *dual* and *plural number* in *adjectives*.

(ii) ***Variant of number-conditioned case syncretism <ins>in personal pronouns</ins>***

    a. The expressions of *accusative*, *genitive*, and *locative case* are identical in the environment of *plural number*;

    b. The expressions of *accusative* and *genitive case* are identical in the environment of *dual number*.

In order to apply this idea to the Slovenian syncretism data we need to establish the feature decomposition of the relevant grammatical categories, starting with number and case. To characterize the singular/plural/dual number opposition, I adopt the feature system in (7) (Conklin 1962; Noyer 1992; Harbour 2008; i.a.): [±singular] distinguishes singular from nonsingular, while [±augmented] distinguishes plural from dual (the [+singular,+augmented] combination is impossible as it is contradictory; see Harbour 2008: 112–4).

(7)

	SG	PL	DU	impossible
singular	+	–	–	+
augmented	–	+	–	+

There exist several versions of case feature systems that vary in subtle ways (see Jakobson 1984a; Bierwisch 1967; Franks 1995; Calabrese 1998, 2008; i.a.). Here I will adopt the version from Calabrese (2008), which is based on a larger sample of languages and designed to capture more case contrasts. In (8) I represent the portion of the feature system relevant for Slovenian.

(8)

	NOM	ACC	GEN	LOC	DAT	INS
peripheral	–	–	+	+	+	+
source	–	–	+	–	–	–
location	–	–	–	+	–	+
motion	–	+	–	–	+	+

Although this is not uncontroversial, I further assume that nominative is the realization of the absence of case features (Jakobson 1984a; Andrews 1982; Falk 1991; Bittner and Hale 1996; Neeleman and Weerman 1998; i.a.), which mainly serves to simplify the characterization of some syncretisms. I mark the presence of case features with K, with the feature values in parentheses.

We can now tackle characterizing the plural/dual syncretism in nouns and adjectives in feature terms. What differentiates nouns and adjectives from personal pronouns is that only the latter can express person contrasts. I encode this with the [±person] feature (more on this below), which allows us to characterize the relevant syncretism pattern as in (9).

(9) ***Case-conditioned number syncretism (feature-based):***
 The expression of [–augmented] is identical to the expression of [+augmented] in the environment of [–person, K(–motion,+peripheral)].

To unpack the statement in (9), the feature that distinguishes plural from dual ([±augmented]) is neutralized (rendered inert) in the environment of nouns and adjectives ([–person]) with genitive and locative case – which are uniquely characterized by [K(–motion,+peripheral)].

The same treatment of the dative/instrumental syncretism with dual number is provided in (10).

(10) ***Number-conditioned case syncretism (feature-based):***
    The expression of [K(–location)] is identical to the expression of [K(+location)] in the environment of [–singular,–augmented, K(+peripheral,+motion)].

The feature distinguishing locative and instrumental from all other cases ([±location]) is neutralized in the environment of dual ([–singular,–augmented]) forms with dative and instrumental case ([K(+peripheral,+motion)]). Essentially, only the case features that uniquely pick out dative and instrumental can be expressed, while the feature that distinguishes them is rendered inert.

Recall also that this is where speakers vary regarding the pattern in personal pronouns. In order to capture the alternative syncretism pattern, we only need an additional statement on top of (10). This statement, given in (11), is just a slightly revised version of (10).

(11) ***Number-conditioned case syncretism – some speakers (feature-based):***
    The expressions of different value combinations of [K(±location,±motion)] are identical in the environment of [+person, –singular,–augmented, K(+peripheral,–source)].

Apart from picking out only personal pronouns ([+person]), what is changed is that more case features are neutralized: the feature distinguishing locative from dative as well as the feature distinguishing both from instrumental ([K(±location,±motion)]). This leaves only the case features that uniquely characterize locative, dative, and instrumental expressible.

Finally, we can re-examine the number and case-conditioned gender syncretism. I adopt the gender feature system in (12) (see Bierwisch 1967; Despić 2017; i.a.), although since we are dealing with total gender syncretism the feature system itself is not essential.[15]

---

15 In nouns, inflection is further determined based on inflection class (see Section 2.3.1). Class distinctions are completely neutralized along with gender in nonsingular locative, dative and instrumental forms. The class syncretism can be modeled in feature terms the same way as the gender syncretism: a complete neutralization of distinctions (see e.g. Müller 2004 on the feature decomposition of inflection classes in relation to Russian nominal inflection).

(12)

	M	**F**	**N**	**impossible**
masculine	+	–	–	+
feminine	–	+	–	+

What is essential here is that the amount of syncretism decreases as we go from personal pronouns to adjectives and finally to nouns – specifically, the amount of cases in which the syncretism is observed decreases. The three patterns can be characterized in feature terms as in (13).

(13) ***Number-conditioned gender syncretism (feature-based)*:**
   The expressions of different value combinations of [±feminine,±masculine] are
   identical:
   a. in the environment of [–singular, K]   (PERSONAL PRONOUNS)
   b. in the environment of
      [–person, –singular, K(+peripheral)]   (ADJECTIVES)
   c. in the environment of
      [–person,+noun,(–verb), –singular, K(+peripheral,–source)]   (NOUNS)

Since the syncretism occurs in both nonsingular numbers, the only number specification needed in the conditioning environment is [–singular]. The case specifications crucially differ: (i) with the personal pronoun pattern in (13a), the presence of an unspecified K in the conditioning environment means that the syncretism is present in all five oblique cases (recall that nominative is characterized by the absence of K); (ii) with the adjectival pattern in (13b), [K(+peripheral)] uniquely picks out the four relevant cases: genitive, locative, dative, and instrumental; and (iii) with the nominal pattern in (13c), [K(+peripheral,–source)] uniquely picks out the three relevant cases: locative, dative, and instrumental. The case specification increases in complexity as the syncretic forms decrease. More importantly, the case specifications are in a subset-superset relation, which is also crucial factor when considering the features characterizing the syntactic categories.

   The syntactic categories are mostly underspecified in (13). Only nouns are fully specified (as [–person,+noun,(–verb)]; the need for [–verb] depends on whether one assumes the Chomsky 1970 analysis, where [±verb] distinguishes nouns from adjectives, since they are both [+noun]). In contrast, the specification for adjectives only involves the absence of person and the syntactic category specification is completely absent with personal pronouns. At first glance this suggests that (13a) should apply to all nonsingular oblique elements (including adjectives and nouns), and similarly that (13b) should also apply to nouns. Note, however, that if the distribution of syncretic forms is governed by the *Paninian/Elsewhere Principle* (Anderson 1969, 1986; Kiparsky 1973; Aronoff 1976; i.a.): more specific forms/rules are preferred over more general ones, there will be no over-application of (13a) and (13b).

Tab. 8: Relative/Wh-pronouns ('which/who').

		NOM	ACC	GEN	LOC	DAT	INS
SG	M.AN	kater**i**	kater**ega**		kater**em**	kater**emu**	kater**im**
	M						
	N	kater**o**					
	F	kater**a**		kater**e**	kater**i**		kater**o**
PL	M	kater**i**	kater**e**	kater**ih**		kater**im**	kater**imi**
	N	kater**a**					
	F						
DU	M	kater**a**				kater**ima**	
	F/N	kater**i**					

All three environment specifications in (13) are in a subset-superset relation with each other and thus governed by the principle, which ensures that more specific syncretic forms are not overrun by more general ones.[16]

The [–person] specification of nouns and adjectives is crucial for (13) to work, which raises the question why the absence of person should be so relevant for determining inflection patterns. There is some good evidence that [–person] describes the natural class of elements that require the adjectival inflection pattern better than any syntactic category features. Namely, what counts as an "adjective" based on inflection is a grab bag in Slovenian (cf. Toporišič 2000: 317–44): adjectives, numerals, quantifiers, possessive nouns and pronouns, demonstrative pronouns, relative/wh-pronouns, etc.; the adjectival inflection pattern of the latter is shown in Table 8 (*kateri* is the inanimate form and *katerega* the animate form with M.SG.ACC; see Section 3.1). The list includes everything from actual adjectives to nouns to closed class elements – what unifies them is that they may all modify nouns and that they never express person oppositions inflectionally.[17]

---

16 Consider with this in mind the ACC/GEN/LOC syncretism in personal pronouns (see footnote 14), where there is no unique feature characterization for the cases in question. The solution is to underspecify the relevant suffixes: **-ih** [–singular] and **-ju** [+person, –singular,–augmented, K] (cf. Table 5). Note that the specification of **-ih** essentially makes it the default nonsingular form. This is intentional, as its distribution across the three inflection patterns is not uniform in terms of case (ACC/GEN/LOC in personal pronouns, GEN/LOC in adjectives, LOC in nouns) or number (plural in personal pronouns, nonsingular in adjectives and nouns). This suggests that its distribution is determined via blocking by more specific forms, which in turn means that the syncretisms associated with **-ih** (**-Vh** on nouns) arise because the form is underspecified rather than, for example, because of the deletion of features (see footnote 18).

17 Possessive pronouns are the reason why the class is defined in terms of inflectional person marking rather than person in general. The stems of possessive pronouns are identical to those of oblique

What is significant about the feature-based descriptions of the syncretism patterns is that they correctly encode the natural class information despite the rather complex interactions between number and other inflectional categories. Additionally, the descriptions accomplish this task with feature systems that were not posited specifically to deal with the Slovenian data.

Coming back to the role of number in all of this consider that: (i) nonsingular numbers are involved in more cases of syncretism than singular number, and (ii) dual is involved in more cases of syncretism than plural. Considering one of the diagnostics for morphosyntactic markedness: either more or the same amount of syncretism in the marked value of a grammatical opposition (Croft 2003: 95–9; see also Greenberg 1966: Ch. 3; Nevins 2011), the described state of affairs conforms to the markedness hierarchy in (14). This can be further decomposed into the markedness statements for individual number features, as in (14a) and (14b) (from Nevins 2011: 421).

(14)   ***Markedness hierarchy:*** dual > plural > singular
      a. ***Context-free markedness statement (feature-based):***
          The marked value of [±singular] is −.
      b. ***Context-sensitive markedness statement (feature-based):***
          In the context [−singular], the marked value of [±augmented] is −.

In fact, Nevins (2011) argues that number syncretism and number-conditioned syncretism are both determined by the markedness of the number value, which he also bases on a comparison of syncretism patterns in standard and Ljubljana Slovenian. The specific syncretisms he considers occur in nominative nouns. In standard Slovenian, neuter and feminine are syncretic in the dual (see Table 9). Conversely, in Ljubljana Slovenian, masculine and neuter are syncretic in the dual, but also plural and dual are syncretic in neuter and masculine forms (see Table 10).

Tab. 9: Standard Slovenian.

	M ('chairs')	N ('windows')	F ('books')
PL	stol**i**	okn**a**	knjig**e**
DU	stol**a**	okn**i**	knjig**i**

Tab. 10: Ljubljana Slovenian.

	M ('chairs')	N ('windows')	F ('books')
PL	stol**i**	okn**a**	knjig**e**
DU	stol**a**	okn**a**	knjig**e**

---

personal pronouns (e.g. *m-* 'me-' ∼ *m-oj* 'me-POSS = my'), which supplete for number and person (see Table 4 in Section 2.2.1). Possessive pronouns thus have an inherent person value, but also inflect based on the features of the nominal possessum, which is always [−person]; see Section 3.1 (e.g. *m-oj-e pesm-i* 'me-POSS-F.PL song-F.PL'). Outside 1st and 2nd person singular forms, possessive pronouns even have two inflectional suffixes: one below the POSS morpheme for the inherent features of the pronoun and one above it for the features of the possessum (e.g. *nj-e-n-emu* 'her-SG.GEN-POSS-M.SG.DAT'). Crucially, only the outer inflection is [−person] and follows the adjectival inflection pattern.

Tab. 11: 1st person pronouns.

	M	**F**
SG	jaz	
PL	m**i**	
DU	m**i**dv**a**	m**i**dv**e**

Tab. 12: 1st person pronouns (variant).

	M	**F**
SG	jaz	
PL	m**i**	m**e**
DU	m**i**dv**a**	m**e**dv**e**

Without going into the details of Nevins' analysis (see his paper for those),[18] we see that the number environment for the gender syncretism in standard Slovenian is [−singular,−augmented], where both feature values are marked. Although the gender syncretism is different in the Ljubljana dialect, the number environment for it is the same. On the other hand, the number syncretism in the Ljubljana dialect neutralizes the distinction between the more marked dual ([−augmented]) and less marked plural ([+augmented]), which causes the less marked form to take over. Note that this characterization is also consistent with the syncretism patterns discussed above, such as the number-conditioned dative/instrumental syncretism (conditioned by [−singular,−augmented]) and the case-conditioned plural/dual syncretism (neutralizing [±augmented]).

Nonetheless, there is a syncretism in Slovenian that goes against the markedness generalization in question. Recall from Section 2.2 that 1st and 2nd person nominative pronouns come in two variants with distinct gender marking patterns – more precisely, two different patterns of gender syncretism. The first pattern, illustrated in Table 11 for 1st person, has total gender syncretism with singular and plural number, while the second pattern, illustrated in Table 12, has total gender syncretism only with singular number (2nd person forms behave the same way; see Section 2.2).

Crucially, in both of the patterns, it is the top of the markedness scale in (14) that does not show syncretism. What we get here is what seems like a reversal of the markedness scale when it comes to syncretism (there is either more or the same amount of syncretism in the marked value of a grammatical opposition), suggesting the alternative scale in (15).

(15)   singular > plural > dual

---

18 Nevins (2011) derives syncretism via morphological *impoverishment* operations which delete the features responsible for the relevant opposition (see Bonet 1991; Noyer 1992; Halle and Marantz 1993; Halle 1997; i.a.). An alternative to this is underspecification of the vocabulary items realizing the features (see Calabrese 1998, 2008; i.a.). See also footnote 16 for some support in favor of the latter approach from some syncretic forms in Slovenian.

In a markedness based analysis, the alternative gender syncretism pattern must be attributed to a *markedness reversal* (Jakobson 1960; Greenberg 1966: 24; Tiersma 1982; Battistella 1996: 58–60; Croft 2003: 164–5, 189–92; also known as *contextual/local markedness*). In other words, the marked values off an opposition become the unmarked ones and vice versa. However, if what we see here really is a markedness reversal, why does it only affect syncretism? There is no evidence that any other markedness-related pattern is affected in the context in question. For example, we can see that dual forms are built off plural ones: *midva* 'we.DU' ~ *mi* 'we.PL', which points to dual being more marked (see Croft 2003: 88–94 and Section 2.3.2). Whether or not the markedness based analysis is on the right track here, it is clear that we should not draw any conclusions on the nature of syncretism or markedness by looking at either nouns or personal pronouns in isolation.

2.2.3 Weak pronouns and number-restricted paradigms

In addition to strong personal pronouns, Slovenian has two types of clitic/weak personal pronouns: (i) second position clitic pronouns, and (ii) weak pronouns that occur only in prepositional phrases, which I call *preposition pronouns*. The latter type are sometimes considered special variants of the regular clitic forms (Herrity 2000: 92–3, 2016: 129–30), but we will see that the two types of weak pronouns behave very differently when it comes to different number values.

The paradigm of second position clitic pronouns is case-restricted: they only exist in accusative, genitive, and dative forms, and bear the same inflection as the corresponding strong pronouns; see Table 13 (nonsingular pronouns are identical to their strong counterparts apart from their syntactic placement and the inability to bear stress). Preposition pronouns, however, exist only in accusative forms and their paradigm is further restricted by person and number value: there are no 1st/2nd person nonsingular preposition pronouns; see Table 14 (the preposition used here is *na* 'on'; the preposition always bears stress when combined with the pronoun).[19] Note that unlike clitic pronouns all 3rd person preposition pronouns contain a support morpheme (underlined; see also Section 2.2).

Crucially, the deficient paradigm of preposition pronouns does not involve syncretism – the nonsingular 3rd person forms do not extend to 1st and 2nd person, as they would with syncretism; in the relevant nonsingular contexts, a strong pronoun form must be used instead.[20]

---

19 In some cases, like *méd.nje* 'between them', the singular form **méd.enj* 'between him' is blocked for semantic reasons – a strong pronoun is also incompatible with the preposition, which requires a distributive reading.

20 A more detailed discussion of the implications of the limited paradigms in contrast to syncretism is impossible here for space reasons, but see Stegovec (2020) for a preliminary analysis.

Tab. 13: Clitic pronouns.

		ACC	GEN	DAT
SG	1	m**e**		m**i**
	2	t**e**		t**i**
	3.M/N	**ga**		**mu**
	3.F	**jo**	**je**	**ji**
PL	1	na**s**		na**m**
	2	va**s**		va**m**
	3	**jih**		**jim**
DU	1	na**ju**		na**ma**
	2	va**ju**		va**ma**
	3	**ju**		**jima**

Tab. 14: Preposition pronouns.

		na+ACC
SG	1	ná.m**e**
	2	ná.t**e**
	3.M/N	ná.<u>nj</u>
	3.F	ná.<u>nj</u>**o**
PL	1	
	2	
	3	ná.<u>nj</u>**e**
DU	1	
	2	
	3	ná.<u>nj</u>**u**

2.2.4 Number-invariant pronouns

Some types of pronouns in Slovenian inflect for case and sometimes animacy, but never for person, gender and number. This includes reflexive pronouns (strong, clitic, and preposition versions) and a variety of pronouns based off the interrogative/indefinite pronoun stem; see Table 15.[21]

The systematic nature of the number-invariance suggests the lack of an inherent number value, although as part of a more general deficiency which also affects gender and person.[22]

---

21 Note that with free relative and negative pronouns, the derivational morpheme *-(V)r* appears outside the inflectional suffix – a counterexample to Anderson's (1982) prediction that *"if a morphologically complex form contains both derivational and inflectional suffixes, the inflectional ones will follow the derivational ones"* (Anderson 1982: 609).

22 While they do not express person inflectionally, reflexive pronouns do pattern with 1st and 2nd person pronouns based on the form and position of the support morpheme (see also Stegovec 2016: 302), which could be attributed to the presence of person features; see Wiltschko (2014: 210–8) regarding a similar situation in German.

Tab. 15: Number-neutral pronouns (non-exhaustive list).

	NOM	**ACC**	**GEN**	**LOC**	**DAT**	**INS**
strong reflexive		seb**e**		seb**i**		sab**o**
clitic reflexive		s**e**			s**i**	
preposition reflexive (+*na* 'on')		ná.s**e**				
interrogative (AN)	kdo	ko**ga**		ko**mu**		ko**m**
interrogative	kaj		čes**a**	č**em**	č**emu**	č**im**
indefinite (AN)	nekdo	neko**ga**		neko**mu**		neko**m**
indefinite	nekaj		nečes**a**	neč**em**	neč**emu**	neč**im**
free relative (AN)	kdor	ko**gar**		komu**r**		kome**r**
free relative	kar		čes**ar**	čem**er**	čem**ur**	č**im**er
negative (AN)	nihče	niko**gar**		niko**mur**		niko**mer**
negative	nič		ničes**ar**	niče**mer**	niče**mur**	nič**im**er

2.3 Nominal number

Nouns express the three-way number opposition in all cases except genitive and locative, although there are minor differences in number marking across inflection classes (Sections 2.3.1). Number also plays a key role in conditioning morphological alternations on the nominal stem, including a rather challenging stem suppletion pattern (Section 2.3.2). Finally, the Slovenian number system interacts in interesting ways with number-deficient nouns (Section 2.3.3) and derived nouns (Section 2.3.4), which include nonsingular event nominalizations that have been argued to not be possible.

2.3.1 Inflection classes and syncretism

Slovenian nominal inflection is largely determined based on the grammatical gender of the noun, although there is some dissociation between gender and the inflection patterns: nouns of different genders may share an inflection pattern and different inflection patterns may arise with nouns of the same gender. Based on this, nouns fall into the four inflection classes illustrated in Table 16;[23] the extra split in

---

23 I follow here the classes from Marušič and Žaucer (2021). Traditional grammars (cf. Toporišič 2000: 276–301) divide nouns of each gender into four inflection classes: the 1st class is the default one, the 2nd class consists of "exceptional" nouns: masculine nouns with and feminine nouns without a vocalic suffix in nominative singular – see below (this class is empty in neuter nouns),

Tab. 16: Nominal inflection classes.

		NOM	ACC	GEN	LOC	DAT	INS
Class I 'sign(M)'	SG	znak		znak**a**	znak**u**		znak**om**
	PL	znak**i**	znak**e**	znak**ov**	znak**ih**	znak**om**	znak**i**
	DU	znak**a**				znak**oma**	
Class I 'bull(M.AN)'	SG	bik	bik**a**		bik**u**		bik**om**
	PL	bik**i**	bik**e**	bik**ov**	bik**ih**	bik**om**	bik**i**
	DU	bik**a**				bik**oma**	
Class II 'house(F)'	SG	hiš**a**	hiš**o**	hiš**e**	hiš**i**		hiš**o**
	PL	hiš**e**		hiš	hiš**ah**	hiš**am**	hiš**ami**
	DU	hiš**i**				hiš**ama**	
Class III 'thing(F)'	SG	stvar		stvar**í**	stvár**i**		stvar**jó**
	PL				stvar**éh**	stvar**ém**	stvar**mí**
	DU					stvar**éma**	
Class IV 'town(N)'	SG	mest**o**		mest**a**	mest**u**		mest**om**
	PL	mest**a**		mest	mest**ih**	mest**om**	mest**i**
	DU	mest**i**				mest**oma**	

Class I is due to the sensitivity to animacy with masculine singular accusative forms.[24] Inflectional suffixes will be segmented out in Slovenian examples from now on in order to make comparison between them easier. Additionally, the inherent grammatical gender of nouns will be indicated in parentheses in order to differentiate it from the contextual gender on agreeing words.

Class I contains masculine nouns, Class II contains feminine and masculine nouns that have an **-a** suffix in nominative singular, Class III contains feminine nouns without the **-a** suffix (and *ljudje* 'people(M)'; see Section 2.3.2), and Class IV contains neuter nouns. Some Class I nouns, like *kmet-**i**/**-je** 'farmer(M)-PL.NOM', have two plural nominative suffix forms in free variation (although it is possible that

the 3rd class consists of nouns that do not inflect, and the 4th class consists of nouns that inflect like adjectives. Nouns in the 4th class actually behave like substantivized adjectives and adjectives modifying a null/elided noun according to the tests of Giannakidou and Stavrou (1999), and will not be considered.

24 Morphological animacy does not always correspond to semantic animacy. Masculine inanimate nouns that inflect like animate ones, which are mostly proper nouns, include among others: car makes and models, diseases, sports teams, wine varieties, devices named after animate entities, card suits and faces, and planets (Toporišič 2000: 277).

a closer inspection could reveal contextual restrictions on the suffix choice). This possibility is more common with animate nouns – in fact, *-je* is the only suffix with some animate nouns (e.g. *mož-je* 'man(M)-PL.NOM'). With the inanimate nouns that also have two plural forms (e.g. *vozóv-i*/*voz-jé* 'wagon(M)-PL.NOM'; note also the stem change, discussed in Section 2.3.2), the *-je* suffix is prescriptively enforced (Toporišič 2000: 283).

Finally, some nouns do not inflect (e.g. some acronyms and abbreviations, some loanwords and foreign names, and terms of endearment like *mami* 'mommy(F)' or *babi* 'granny(F)'). As these nouns cannot express number inflectionally, their number is revealed only via agreement.

Recall from Section 2.2.2 that nouns and adjectives systematically show plural/ dual syncretism in genitive and locative case. Class III nouns show the same plural/ dual syncretism, but as part of a more general syncretism pattern: the same form (*stvar-í* in Table 16) is used in nonsingular nominative and accusative, as well for all three numbers in genitive case. Certain Class III nouns, like *miš* 'mouse(F)', lack the variable stress placement, which means that they display a further syncretism with locative/dative singular (*míš-i* 'mouse(F)-SG.GEN' = 'mouse(F)-SG.LOC/DAT'), although this is most likely an accidental syncretism due to it resulting from a phonological factor.[25]

2.3.2 Stem changes and suppletion

Unequivocal cases of number-based suppletion in nouns are very rare in Slovenian. Most number-sensitive irregularities in noun stems can be characterized as morphologically conditioned phonological alternations like stress-shift (e.g. *móž* 'man(M).SG' ∼ *mož-jé* 'man(M)-PL'[26]) or regular morphological processes like adding to the stem a morpheme that does not by itself express an inflectional property – an *augment*. For example, in (16) an augment occurs with nonsingular forms (although some speakers restrict the augment to plural forms – the dual form is then *móst-a*).[27]

---

25 In the few dialects that retain a pitch accent system (Toporišič 2000: 63–4; Herrity 2000: 4, 10–1, 2016: 12, 27–8), there is less syncretism with variable stress Class III nouns: *-î* is the NSG.NOM/ ACC and SG.GEN suffix (most likely accidental syncretism), while *-ī* is the NSG.GEN suffix. Note that this pattern maintains the plural/dual syncretism in genitive and locative case, and in addition shows nominative/accusative syncretism with plural and dual number.
26 For convenience sake, the case information is left out in the glosses of nominative (citation) forms throughout the remainder of the chapter, unless the contrast between nominative and another case is of interest.
27 The availability of the alternative plural-only pattern also appears to be subject to lexical variation. For example, a preliminary search in the *Gigafida 2.0* corpus of written standard Slovenian (https://viri.cjvt.si/gigafida/; accessed June 2020), reveals that the nominative dual form of *volk* 'wolf(M)' is augmentless in 33 of the 93 entries, while the nominative dual form of *grad* 'castle(M)' is augmented in all 144 entries.

(16) *móst* 'bridge(M).SG' ~ *most-óv-i* 'bridge(M)-AUG-PL' ~
      most-óv-a 'bridge(M)-AUG-DU'

Importantly, augmentation is not only conditioned by number, as it is restricted to select nouns of specific inflection classes and by phonological properties of the stem; for example, the *-ov* augment in (16) is restricted to select monosyllabic Class I nouns. Furthermore, with the Class I nouns where two nominative plural suffixes are available, the choice of suffix may also condition stem augmentation, as with: *voz-óv-i* 'wagon(M)-AUG-PL₁' versus *voz-jé* 'wagon(M)-PL₂'. The augment is also not always semantically vacuous. In some nouns where the augment is prescribed against (Toporišič 2000: 283), the colloquial augmented variant can develop a specialized meaning. For example, unlike the augment-less plural of 'tooth' (*zób-i*), the augmented version is becoming specialized to mean protruding parts of a tool, instrument, or other object, as illustrated in (17).[28]

(17) a. *jermen   z   vgrajenimi kovinski-mi zob-óv-i*
        transmission.belt with built.in   metal-PL.INS tooth(M)-AUG-PL.INS

   b. #*z moj-imi zob-óv-i* 'with my-M.PL.INS tooth(M)-AUG-PL.INS'

Similarly to stem augmentation, other number-sensitive changes on the stem are not just sensitive to number. For example, palatalization of the stem final consonant in the plural forms of *otrok* 'child(M)' is also conditioned by case (Priestly 1993: 401; Toporišič 2000: 283; Herrity 2000: 50, 2016: 73). Thus, while the *k→c* palatalization only occurs with plural forms (cf. (18a)), only nominative and locative plural stems show it (cf. (18b)); the pattern also shows the alternation cannot be purely phonological, as the instrumental suffix is phonologically identical to the nominative one.

(18) a. *otrok* 'child(M).SG' ~ *otrok-a* 'child(M)-DU' ~ *otroc-i* 'child(M)-PL'

   b. *otroc-i* 'child(M)-PL.NOM' ~ *otroc-ih* 'child(M)-PL.LOC' ~
      otrok-i 'child(M)-PL.INS'

The conditions on stem modifications are thus quite complex and not exclusively sensitive to number. The situation is similar with the one clear example of stem suppletion. The example in question, the alternation between *človek-* 'person' and *ljud-* 'people', has been discussed extensively (see Plank 1994; Corbett 2000: 275–6; Evans, Brown, and Corbett 2001: 215; Baerman, Brown, and Corbett 2005: 175–7; Smith et al. 2019: 1075; i.a.). The suppletion pattern in question is shown in Table 17.

    While the singular stem is always *človek-* and the plural stem is always *ljud-*, the situation is more complicated with dual forms. In nominative, accusative, da-

---

28 My intuitions are confirmed by data from the *Gigafida 2.0* corpus (https://viri.cjvt.si/gigafida/; accessed June 2020), where all 5 augmented entries of 'teeth' conform to the generalization.

Tab. 17: Number stem suppletion pattern: *človek-* ~ *ljud-* ('person' ~ 'people').

		NOM	ACC	GEN	LOC	DAT	INS
Class I.AN	SG	člóvek	človeéka		človeéku		človeékom
"person"	PL	ljudjé	ljudí		ljudéh	ljudém	ljudmí
	DU	človeéka			človeékoma		

tive, and instrumental case the dual stem is identical to the singular stem, whereas in genitive and locative case the dual stem is identical to the suppletive plural stem, so *ljud-* is neither exclusivelly the plural nor exclusively the nonsingular stem. Note also that *človek-* stems inflect like Class I masculine animate nouns, but *ljud-* stems inflect, apart from the nominative plural suffix, like Class III feminine nouns.[29]

Recall that genitive and locative are also cases where plural/dual syncretism occurs in all nouns, which is something that all existing analyses of the pattern make reference to, although in different ways. For example, Evans, Brown, and Corbett (2001) invoke *rules of referral* (Zwicky 1985; Stump 1993), which specify that one cell in the paradigm is realized identically to another – in this case, plural and dual forms in genitive and locative case. The nonsingular genitive and locative forms thus share a realization across the nominal paradigm, meaning both the inflection and suppletion pattern, and the suppletive *ljud-* stem is always plural. Alternatively, Smith et al. (2019) adopt a markedness-based approach and characterize the pattern as suppletion driven by a marked plural (Smith et al. 2019: 1074); in feature terms, [+augmented] is a marked value. For the exceptional genitive and locative cases, Smith et al. suggest that these simply lack a dual number and therefore dual genitive and locative forms (Smith et al. 2019: 1075); the syncretism is due to underspecification rather than rules of referral. Given the feature system the authors assume (see (7) in Section 2.2.2) this implies that only with genitive and locative case [–singular] drives suppletion – suppletion always occurs with plural forms, but the feature driving it is not always the same one.[30] Another complication concerns the assumption that plural is marked and dual unmarked in Slovenian. Recall from Section 2.2.2 that dual behaves as marked with respect to syncretism in nouns. Furthermore, nominative dual personal pronouns contain their plural forms (cf. *m-i* 'we-PL' ~ *m-i-dv-a*

---

29 To the extent that we can draw conclusions from such an irregular noun, the co-existence of multiple inflection patterns with one lexeme, or *heteroclisis*, can be taken as evidence that inflection class is a property of stems rather than lexemes; see Stump (2002, 2006) for discussion.

30 A related analysis that avoids this is sketched out by Bobaljik (2002: 81n14): the plural/dual syncretism in genitive and locative results from the impoverishment of [–augmented] followed by the insertion of the unmarked [+augmented] value (cf. *persistent redundancy rules* of Noyer 1998). This way, both dual and plural forms in genitive and locative are specified as [+augmented] and the suppletive *ljud-* stem is always associated with the same feature.

'we-PL-two-DU'), which Smith et al. also associate with dual being more marked than plural (Smith et al. 2019: 1073).[31] The marked plural analysis therefore does not match up with data outside this case of suppletion.

Something that is ignored by all analyses of the *človek- ∼ ljud-* suppletion that I am aware of is that it is also attested in cases where there is no apparent number opposition in play. We see this with compounds, as in (19a) (LK = linking morpheme), and with derivational processes, as in (19b).

(19) a. *človek-o-ljub*   vs.   *ljud-o-žer*
          person-LK-love   people-LK-devour
          'philanthropist(M)'   'cannibal(M)'

     b. *človeš-tvo*   vs.   *ljud-stvo*
          person-DER   people-DER
          'humanity(N)'   'the people(N)'

Although the words formed with the different stems differ in meaning, the difference is in no way transparently related to number. Furthermore, when we look at inflectional number contrasts on such derived words, we see that they do not affect the suppletion pattern at all:

(20) *človeš-k-a/-e*   vs.   *ljud-sk-a/-e*
       person-DER-F.SG/-F.PL   people-DER-F.SG/-F.PL
       'human (adjective)'   'folk/populist (adjective)'

One way to view the derivational suppletion pattern is as entirely unrelated to the inflectional one we saw above – that is, number is not the only trigger for the *človek- ∼ ljud-* suppletion. Under this view, it becomes somewhat surprising that the suppletive stem is not more common. We are also left with the question of what is the suppletion trigger in the relevant derivational contexts.

An alternative possibility is to relate the alternation to "plural" linking morphemes in Germanic, as in Dutch *schaap-en* 'sheep-PL' versus *schaap-en-bout* 'sheep-LK-leg' (Booij 2001: 179; Krott, Schreuder, and Baayen 2002; Krott et al. 2007; De Belder 2013; Fenger and Harðarson 2018, 2020). These are homophonous with plural morphology and also do not necessarily convey a plural meaning. De Belder (2013) and Fenger and Harðarson (2018, 2020) propose that such linking morphemes are derivational morphemes with a number specification (see also Acquaviva 2008;

---

31 Dual forms also contain plural forms with dative personal pronouns, nouns, and adjectives (e.g. *na-m* 'us-PL.DAT' ∼ *na-m-a* 'us-PL.DAT-DU', *rak-om* 'crab(M)-PL.DAT' ∼ *rak-om-a* 'crab(M)-PL.DAT-DU', *rdeč-im* 'red-PL.DAT' ∼ *rdeč-im-a* 'red-PL.DAT-DU'), but the decomposition is not as clean when morphemes express both number and case.

Kramer 2016 for related cases), where the plural meaning of the feature can remain semantically uninterpreted (Sauerland 2003; Sauerland, Anderssen, and Yatsushiro 2005). A way to reinterpret this for the cases in (19)–(20) would be to treat the suppletive *ljud-* stem as a derivational plural form. The reason why inflectional number cannot override the choice of stem could then be attributed to locality restrictions on suppletion triggers (Moskal 2015a,b): the presence of derivational morphemes prevents inflectional number – which is further away from the stem – to act as the conditioning environment for stem suppletion.

If this analysis sketch is on the right track, then we hope to find additional cases of derivational number morphemes in Slovenian. A possible candidate for this kind of morpheme are the stem augments (like *-ov*) discussed above; their distribution is sensitive to number and is crucially lexically idiosyncratic. Furthermore, as we will see briefly in Section 2.3.4, the appearance of augments is sometimes conditioned not by inflectional number but by derivational processes.

2.3.3 Lexical irregularities in inherent number

Although nouns are inherently specified for number, some are number deficient: they are limited to a specific number, but not because of an inflectional deficiency. The basic example of number deficient nouns are *pluralia tantum*, which exist only in plural form regardless of the true number of the referent(s) (if the noun is also a plurale tantum in English, the gloss is plural):

(21) ***Pluralia tantum:***
*anal-**i*** 'annals(M)-PL', *možgan-**i*** 'brain(M)-PL', *obrest-**i*** 'interest(M)-PL (finance)', *citr-**e*** 'zither(F)-PL', *hlač-**e*** 'trousers(F)-PL', *rovt-**e*** 'backwater(F)-PL (pej.)', *škarj-**e*** 'scissors(F)-PL', *vic-**e*** 'purgatory(F)-PL', *jetr-**a*** 'liver(N)-PL', *nebes-**a*** 'heaven(N)-PL', *očal-**a*** 'glasses(N)-PL', *tl-**a*** 'ground/floor(N)-PL', *ust-**a*** 'mouth(N)-PL', *vrat-**a*** 'door(N)-PL', ...

In Slovenian, pluralia tantum inflect and behave syntactically like regular plural nouns; for example, they trigger plural agreement (see Sections 3.1 and 3.2). Some examples like *nebes-**a*** 'heaven(N)-PL' have apparent singular counterparts, in this case *neb-**o*** 'sky(N)-SG'. However, the two have different meanings and do not constitute a productive number alternation – the singular 'sky' is also a mass noun and can only be singular (see below). All canonical pluralia tantum nouns, exemplified in (21), are inanimate. Other than including a large number of body parts, instruments, and abstract concepts, there is no clear cut property unifying all the pluralia tantum in Slovenian.

In addition to pluralia tantum, Slovenian also has *singularia tantum* (see below), but interestingly no *dualia tantum*. It does, however, has a class of nouns that

one would expect to be dual based on their meaning, but which occur in plural form instead (Priestly 1993: 440–1; Toporišič 2000: 271; Herrity 2000: 38–9, 2016: 59–60; Corbett 2000: 93–4; Derganc 2003: 172–4; i.a.):

(22) *Nog-e   me   bolijo.*
 leg(F)-PL 1SG.DAT hurt.3PL
 'My legs hurt.'
 (Derganc 2003: 172)

I refer to these nouns as *pair nouns* since they all refer to entities that prototypically come in pairs:

(23) **Pair nouns:**
 a. paired body parts: *oč-i* 'eye(M)-PL', *kolen-a* 'knee(F)-PL', *nog-e* 'leg(F)-PL', *nosnic-e* 'nostril(F)-PL', *perut-i* 'bird wing(F)-PL', *rok-e* 'hand(F)-PL', *ušes-a* 'ear(N)-PL', ...
 b. paired items of clothing/equipment: *čevlj-i* 'shoe(M)-PL', *rokav-i* 'sleeve(M)-PL', *drsalk-e* 'ice skate(F)-PL', *kotalk-e* 'roller skate(F)-PL', *noga-vic-e* 'sock(F)-PL', ...
 c. paired biological roles: *dvojčk-i* 'twin(M)-PL', *starš-i* 'parent(M)-PL', ...

Crucially, these pair nouns can also appear in dual form. In fact, they must be dual when modified by *dva* 'two' or *oba* 'both', as shown in (24) (some exceptions will be discussed in Section 4.4).

(24) *Ob-e   nog-i   me   bolita.*
 both-F.DU leg(F)-DU 1SG.DAT hurt.3DU
 'Both my legs hurt.'
 (Derganc 2003: 172)

The fact that these nouns can either be plural or dual when referring to two entities is used by Corbett (2000) to argue that dual number is facultative (i.e. optional) in Slovenian nouns. Alternatively, Derganc (2003) argues that plural pair nouns are a kind of pluralia tantum: *nog-e* 'leg(F)-PL' in (22) is a plurale tantum noun – plural in form but not in meaning, while *nog-i* 'leg(F)-DU' in (24) is a regular dual noun. This implies the two forms do not constitute a regular number alternation, just like the *nebes-a* 'heaven(N)-PL' versus *neb-o* 'sky(N)-SG' do not. Furthermore, the absence of animate pluralia tantum nouns noted above should be reinterpreted as a tendency due to the nouns in (23c). Assessing the viability of either analysis requires a closer look at the differences in meaning between the plural and dual uses of pair nouns, which I postpone until Section 4.4.

Turning back to more clear-cut cases, the singularia tantum nouns in Slovenian can be divided into two groups: *mass nouns* and *group/collective nouns*. Some examples are given in (25)–(26).

(25) **Mass nouns:**
      čaj 'tea(M).SG', *denar* 'money(M).SG', *fižol* 'beans(M).SG', *riž* 'rice(M).SG', *zrak* 'air(M).SG', *kr-i* 'blood(F)-SG', *rž* 'rye(F).SG', *svetlob-a* 'light(F)-SG', *trav-a* 'gras(F)-SG', *vod-a* 'water(F)SG', *goriv-o* 'fuel(N)-SG', *neb-o* 'sky(N)-SG', *vin-o* 'wine(N)-SG', *zlat-o* 'gold(N)-SG', ...

(26) **Group/collective nouns:**
      divjad 'game(F).SG (hunting)', *bralstv-o* 'readership(N)-SG', *goved-o* 'cattle(N)-SG', *grmov-je* 'bushes(N)-SG', *grozd-je* 'grapes(N)-SG', *klas-je* 'ears of wheat(N)-SG', *list-je* 'foliage(N)-SG', *peril-o* 'linen/laundry(N)-SG', *skalov-je* 'rocks(N)-SG', *zverjad* 'beasts of pray(N).SG', ...

Mass nouns are inanimate and, as is common, include mainly substances, liquids, and materials. Group/collective nouns are a much more open class that includes many derived nouns (e.g. *bral* 'he read' ~ *bral-stv-o* 'readership(N)-SG', *list* 'leaf(M)' ~ *list-je* 'foliage(N)-SG') (see Section 2.3.4). Unsurprisingly, mass nouns can become countable with the addition of a measure word, as in (27).

(27) a. *dv-e zrn-i riž-a* 'two-F.DU grain(F)-DU rice(M)-SG.GEN'
      b. *dv-a litr-a vod-e* 'two-M.DU liter(M)-DU water(F)-SG.GEN'
      c. *tr-i flaš-e vin-a* 'three-F.PL bottle(F)-PL wine(N)-SG.GEN'

In such cases, the mass noun remains singular and gets genitive case. The measure word is a noun with its own gender and number and the head of the phrase, as evidenced by the fact that it is the measure word rather than the mass noun which controls agreement; e.g. on the numeral in (27). The exact same pattern is also observed with certain quantifiers and numerals (see Section 3.2).

An alternative way to count mass nouns is by way of *recategorization* (see Corbett 2000: 84–7), where a bare mass noun gets either a *type reading*, as in (28a), or a *unit reading*, as in (28b).

(28) a. *Od   točen-ih   so   imeli   **tri**   piv-e.*   (*type*)
         From tap-PL.GEN AUX.3PL had.M.PL three.PL.ACC beers(F)-PL.ACC
         'They had three beers on tap.'

      b. *Častil   nam   je   **tri**   piv-e.*   (*unit*)
         treated.M.SG 1PL.DAT AUX.3SG three.PL.ACC beers(F)-PL.ACC
         'He treated us to three beers/rounds of beers.'

Both readings are possible also with dual forms of mass nouns, so (29) can have either reading.

(29) *Sir-**a**,*   *ki*   *ste*   *nam*   ***ju***   *ponudili,*   *bomo*
     cheese(M)-DU which AUX.2PL 1PL.DAT 3DU.ACC offered.M.PL FUT.1PL
     kupili.
     buy.M.PL
     'The two cheeses (types/units) you have offered us, we will buy.'
     (Corbett 2000: 86–7)

The reverse scenario occurs with some countable nouns, which can get a non-strict general number reading when singular: the singular noun can refer to more than one entity (see also Section 4.4 on other non-strict number readings). Outside generic contexts (see Section 4.2), this reading mainly occurs with: countable foods (cf. (30a)), and animals in fishing/hunting contexts (cf. (30b)).

(30) a. *Jedli*   *bomo*   *puran-**a***   */ potic-**o**.*
       eat.M.PL FUT.1PL turkey(M)-SG.ACC / potica(F)-SG.ACC
       'We'll be eating turkey / potica (nut roll).'

     b. *Lovili*   *bomo*   *som-**a***   */ lisic-**o***   */ #slon-**a**.*
       hunt.M.PL FUT.1PL bass(M)-SG.ACC / fox(F)-SG.ACC / elephant(M)-SG.ACC
       'We'll be hunting/fishing for bass / fox / #elephant.'

Interestingly, the reading is not available in the (30b) scenario if hunting/fishing of the species is not sufficiently conventionalized, as with elephants (although the reading is more easily available if the verb is nominalized: *lov na slon-**a*** 'hunt.M.SG on elephant(M)-SG.GEN' = 'elephant hunting').

2.3.4 Derived nouns

In her seminal work, Grimshaw (1990) observed that not all verb nominalizations can be pluralized: *complex event nominalizations*, which denote events just like their verbal counterparts, cannot be pluralized (e.g. **the destructions of the city*), whereas *result nominalizations*, which instead denote the result state of the verb, arguments of the verb, or concrete objects related to the verb, can be pluralized (e.g. *the reports of my death*). However, in Slovenian even event nominalizations may be pluralized, as shown in (31), although certain restrictions do apply, as we will see.

(31) a. *Zatajil*   *bo*   *pri izvajan-**ju***   *prost-ih*
       mess.up.M.SG FUT.3SG at carrying.out(N)-SG.LOC free-M.PL.GEN
       met-ov.
       throw(M)-PL.GEN

   b. *Zatajil   bo   pr izvajanj-**ih**   prost-ih*
      mess.up.M.SG FUT.3SG at carrying.out(N)-PL.LOC free-M.PL.GEN
      met-ov.
      throw(M)-PL.GEN
      'He will mess up at making free throws.'

There is a difference in meaning between (31a) and (31b): only (31b) infers multiple free throwing sequences (each sequence possibly involving multiple free throws). Crucially, this repetitive reading is not limited to plural themes, as shown in (32): these deverbal nouns crucially denote generic eventualities (most clearly in (32a)) and can either be plural, as in (32a,b), or dual, as in (32c).

(32) a. *Ponovn-**a**   pre-biranj-**a**   dobr-e   knjig-e   so*
      renewed-N.PL PFV-reading(N)-PL good-F.SG.GEN book(F)-SG.GEN AUX.3PL
      priporočena.
      advised.N.PL
      'Repeatedly re-reading a good book is advised.'

    b. *Zakon   mora   v parlament-u   skozi   **tri***
      law(M).SG must.3SG in parliament(M)-SG.LOC through three.PL.ACC
      *branj-**a**.*
      reading(N)-PL.ACC
      'A law must be read three times in parliament.'

    c. *Za sprejet-je   zakon-a   sta   potrebni   še   dv-**e***
      for accept(N)-SG.ACC law(M)-SG.ACC AUX.3DU needed.F.DU also two-N.DU
      *branj-**i**.*
      reading(N)-DU
      'For the law to be accepted, it must be read two more times.'

As Alexiadou, Iordăchioaia, and Soare (2010a,b) note, many counterexamples to Grimshaw's generalization have been found since it was proposed (see references for examples), and these Slovenian examples constitute another one. Alexiadou, Iordăchioaia, and Soare (2010a,b) also argue for an alternative restriction on plural deverbal nouns: *perfective* nominalized verbs can be plural (in our case nonsingular), while *imperfective* ones must be singular. At first glance, (32b) and (32c) above seem to be counterexamples, as the nominalized verbs are in their imperfective form – although the readings suggest completed events. We can control for (im)perfectivity better by using an inherently imperfective verb and the right adverbials. Once we do that, we see that pluralization is restricted:

(33) a. *Pisanj-**e**   pism-**a**   je   trajalo   cel-o*
      writing(N)-SG letter(N)-SG.GEN AUX.3SG lasted.N.SG whole-F.SG.ACC
      noč.
      night(F).SG.ACC

b. ?*Pisanj-a *pism-a* *so* *trajala* *cel-o*
  writing(N)-PL letter(N)-SG.GEN AUX.3PL lasted.N.PL whole-F.SG.ACC
  noč.
  night(F).SG.ACC
  'The writing of the letter took the whole night.'

Note that (33b) is not marked as fully ungrammatical. That is because a grammatical iterative reading can be coerced: there were several sequences of writing of the same letter where no sequence alone resulted in the completion of the letter, or the letter even remained uncompleted. If Alexiadou, Iordăchioaia, and Soare (2010a,b) are correct in positing a link between imperfective verbs and mass nouns on one hand (both are *unbounded*) and count nouns and perfective verbs on the other (both are *bounded*), then the coerced iterative reading could be seen as equivalent to re-categorization with mass nouns (recall (28)–(29) in Section 2.3.3). This type of aspectual flexibility of Slovenian verbs is therefore a promising avenue to explore in relation to the number-aspect connection.

  Substantivized adjectives are another interesting case, although they do not involve transparent morphological nominalization. They are quite productive in Slovenian and inflect the same as regular adjectives (see also footnote 23). Some substantivized adjectives require a specific number value: generic substantivized adjectives referring to humans (cf. (34a)) must be plural and masculine, while those referring to inanimate entities (cf. (34b)) must be singular and neuter.[32]

(34) a. *Bogat-i* *postajajo* *bogatejš-i, revn-i* *pa* *revnejš-i.*
  rich(M)-PL becoming.3PL richer-M.PL poor(M)-PL CNTR poorer-M.PL
  'The rich are getting richer, while the poor are getting poorer.'

b. *Saj* *jem* *tudi sladk-o,*   *ampak zdaj mi* *bolj* *paše*
  PTCL eat.1SG also sweet(N)-SG.ACC but now 1SG.DAT more crave.3SG
  slan-o.
  salty(N)-SG.ACC
  'I do also eat sweet things, but now I'm craving salty things more.'

Finally, derived group/collective nouns can show number-related morphology below the derivational morpheme (*-je*): the stem is nonsingular even though the derived noun is singular. We see this with nouns that have stem augments in nonsingular numbers (see Section 2.3.2), like (35b) and (35c).

---

32 An anonymous reviewer suggests that these can also be plural, e.g. when discussing different dishes in a large buffet: 'The savory ones, the sweet ones ...'. However, at least according to my judgments, these do not have the same generic reading in Slovenian as the substantivized adjectives in (34b) do. This suggests that these plural adjectives might involve an elided noun rather than true substantivization; see Giannakidou and Stavrou (1999) for discussion.

(35) a. *list-**i*** 'leaf(M)-PL' ∼ *list-**je*** 'leaf(N)-DER.SG' = 'foliage'

    b. *rog-<u>ov</u>-**i*** 'horn(M)-AUG-PL' ∼ *rog-<u>ov</u>-**je*** 'horn(N)-AUG-DER.SG' = 'antlers'

    c. *kol-<u>es</u>-**a*** 'wheel(N)-AUG-PL' ∼ *kol-<u>es</u>-**je*** 'wheel(N)-AUG-DER.SG' = 'machine wheels'

Interestingly, the augment sometimes occurs even with nouns that do not show it otherwise:

(36) *skal-**e*** 'rock(F)-PL' ∼ *skal-<u>ov</u>-**je*** 'rock(N)-AUG-SG' = 'rocks/cliffs'

The lexical distribution of the nonsingular augment is thus not idiosyncratic only in relation to inflectional number, its appearance is also further conditioned by derivational processes similarly to suppletion – as we saw previously in Section 2.3.2 with *človek-* ∼ *ljud-* in (19) and (20).

2.4 Verbal number

There is no productive system of verbal number in Slovenian (neither event nor participant number). There are only non-productive cases that can be seen as a kind of participant number effect, where the lexical semantics of the verb is incompatible with a semantically singular argument (comparable to contrasts like 'We/*I gathered outside' in English). Such cases often involve verbs with lexical prefixes, where the prefix alters the aspect, aktionsart, or lexical semantics of the underlying verb stem. For example, with ***u-**sesti se* 'sit down' versus ***po-**sesti se* 'sit down in a sequence/pattern', only the former is compatible with a semantically singular subject, as illustrated in (37)–(38).[33]

(37) a. *Fant   se   je   **u-**sedel.*
      boy(M).SG REFL AUX.3SG sat.down.M.SG
      'The boy sat down.'

    b. *Fant-**je**   so   se   **u-**sedli.*
      boy(M)-PL AUX.3PL REFL sat.down.M.PL
      'The boys sat down.'

(38) a. #*Fant   se   je   **po-**sedel.*
      boy(M).SG REFL AUX.3SG sat.down.M.SG
      'The boy sat down (in a sequence/pattern).'

---

33 The reflexive clitic *se* is required for intransitive use with these two verbs. Also, ***po-**sesti se* may take singular subjects with its alternative meaning 'collapse', mainly limited to inanimate arguments, like buildings.

b. *Fant-**je*** *so* *se* ***po**-sedli.*
boy(M)-PL AUX.3PL REFL sat.down.M.PL
'The boys sat down (in a sequence/pattern).'

There is no transparent correlation between specific lexical prefixes and specific number restrictions. As shown in (39)–(40), there are even cases where the bare verb form has a number restriction (in this case, the object must be plural) and adding the ***po-*** prefix can lift the restriction. These number restrictions are therefore highly idiosyncratic and lexically determined.

(39) a. #*Bral* *sem* *rož-**o**.*
picked.M.SG AUX.3SG flower(F)-SG.ACC
'I was picking a flower.'

b. *Bral* *sem* *rož-**e**.*
picked.M.SG AUX.3SG flower(F)-PL.ACC
'I was picking flowers.'

(40) a. ***Po**-bral* *sem* *rož-**o**.*
picked.up.M.SG AUX.3SG flower(F)-SG.ACC
'I picked up the/a flower.'

b. ***Po**-bral* *sem* *rož-**e**.*
picked.up.M.SG AUX.3SG flower(F)-PL.ACC
'I picked up the/some flowers.'

3 Agreement and the syntax of number

The discussion so far focused on syntactic categories inherently specified for number: pronouns and nouns. The two other main syntactic categories: verbs and adjectives, may acquire number only through agreement. Number is expressed in Slovenian on all agreeing syntactic categories: main and auxiliary verbs (cf. (41)), and both attributive and predicative adjectival forms (cf. (42)).

(41) a. *Želv-**e*** *hodi-**jo** počasi.*
turtle(F)-PL walk-3PL slowly
'Turtles walk slowly.'

b. *Želv-**e*** ***so*** *hodil-**e*** *tam.*
turtle(F)-PL AUX.3PL walked-F.PL there
'Turtles walked there.'

(42) a. *s   počasn-**imi** želv-**ami***
       with slow-PL.INS turtle(F)-PL.INS
       'with slow turtles'

   b. *Želv-**e**   **so**   počasn-**e**.*
       turtle(F)-PL.NOM AUX.3PL slow-F.PL.NOM
       'Turtles are slow.'

In the continuation, I first examine the different patterns of number inflection on nominal modifiers (Section 3.1) and how quantifiers and numerals specifically relate to different number and case marking patterns on the modified noun phrase (Section 3.2). Following that, I consider the different patterns of number inflection on verbs (Section 3.3). We will see that the modifier-conditioned case on the head noun has important ramifications in the domain of agreement (Section 3.4).

3.1 Nominal modifiers and the expression of number

All inflected nominal modifiers follow the adjectival inflection pattern shown in Table 18 (the forms with the *-i* suffix in parentheses are the "long form" adjectives which mark specificity). How different nominal modifiers agree in number with the head noun is illustrated in (43).

(43) ***Adjectives:***

star pes 'old.M.SG dog(M).SG'
*star-**i** ps-**i*** 'old-M.PL dog(M)-PL'
*star-**a** ps-**a*** 'old-M.DU dog(M)-DU'

Numerals:

en pes 'one.M.SG dog(M).SG'
*tri-**je** ps-**i*** 'three-M.PL dog(M)-PL'
*štir-**je** ps-**i*** 'four-M.PL dog(M)-PL'
*dv-**a** ps-**a*** 'two-M.DU dog(M)-DU'

Quantifiers:

vsak pes 'each.M.SG dog(M).SG'
*vs-**i** ps-**i*** 'all-M.PL dog(M)-PL'
*ob-**a** ps-**a*** 'both-M.DU dog(M)-DU'

Possessive pronouns/nouns:

moj pes 'my.M.SG dog(M).SG'
*moj-**i** ps-**i*** 'my-M.PL dog(M)-PL'
*moj-**a** ps-**a*** 'my-M.DU dog(M)-DU'
mam-in pes 'mom-'s-M.SG dog(M).SG'
*mam-in-**i** ps-**i*** 'mom-'s-M.PL dog(M)-PL'
*mam-in-**a** ps-**a*** 'mom-'s-M.DU dog(M)-DU'

Adjectives almost always inflect (there are only a few with zero inflection: e.g. *poceni* 'cheap', *zanič* 'useless', *roza* 'pink', *bež* 'beige', *fajn* 'neat/nice') and possessive pronouns always inflect. Quantifiers and numerals – which I refer to jointly as *quantity modifiers* or QMODs unless disambiguation is needed – are not a homogeneous group regarding inflection. Following Giusti and Leko (1995, 2005), QMODs can be divided into three categories based on their inflection pattern and syntactic behav-

Tab. 18: Adjectival inflection pattern (*star* 'old').

		NOM	ACC	GEN	LOC	DAT	INS
SG	M.AN	star(**i**)	star**ega**		star**em**	star**emu**	star**im**
	M						
	N	star**o**					
	F	star**a**		star**e**	star**i**		star**o**
PL	M	star**i**	star**e**	star**ih**		star**im**	star**imi**
	N	star**a**					
	F						
DU	M	star**a**				star**ima**	
	F/N	star**i**					

Tab. 19: Inflection on quantity modifiers.

				NOM	ACC	GEN	LOC	DAT	INS
Quantity adjectives	'two'	DU	m	dv**a**		dv**eh**		dv**ema**	
			F/N	dv**e**					
	'both'	DU	m	ob**a**		ob**eh**		ob**ema**	
			F/N	ob**e**					
	'three'	PL	m	tr**ije**	tri	tr**eh**		tr**em**	tr**emi**
			F/N						
Quantifiers	'five'	PL	M/F/N	pet		pet(**ih**)		pet(**im**)	pet(**imi**)
	'more'	SG	M/F/N	več					
		PL	M/F/N			več(**ih**)		več(**im**)	več(**imi**)
	'some'	SG/PL	M/F/N	nekaj					

ior: (i) *(true) quantifiers*, (ii) *quantity adjectives*, and (iii) *quantity nouns*. In Slovenian, true quantifiers further divide into optionally inflected ones and uninflected ones.

The differences in the expression of number, gender, and case across QMODs – excluding quantity nouns, which inflect like regular nouns – are illustrated in Table 19 (note the presence of an additional feminine/neuter syncretism with inflected plural numerals, exemplified by 'three').

Quantity adjectives always inflect, although *dva* 'two' and *oba* 'both' are limited to dual forms and the *tri* 'three' to plural ones – presumably for semantic reasons.

Interestingly, the numeral *en* 'one' is not number restricted if the noun phrase it modifies is semantically singular, as with the plurale tantum *vrat-a* 'door(N)-PL' in (44a), or if the numeral denotes one set of nonsingular entities, as in (44b) (see also Section 3.2 regarding the "set reading").[34]

(44) a. *en-**a***   *vrat-**a***
    one-N.PL   door(N)-PL
    'one door'

   b. *En-**e***   *palačink-**e**,*   *prosim!*
    one-F.PL.ACC   crepe(F)-PL.ACC   please
    'One order of crepes, please!'

Note also the differences between the three true quantifiers shown in Table 19. The numeral *pet* 'five' and quantifier *več* 'more' both inflect only optionally, although only 'more' can also be singular in which case it does not inflect. Finally, the quantifier *nekaj* 'some' is always uninflected, regardless of number. As we will see next, the inflection pattern on the quantifier or numeral, more generally as well as in this particular example, also correlates to the case borne by the head noun.

3.2 Case on quantified noun phrases

The choice of QMOD crucially determines the case of the noun phrase (NP) it quantifies over: the entire NP may get external case – that is, whichever case is assigned to it by an external assigner – or the NP excluding the QMOD may get genitive case – a common phenomenon in Slavic languages known as the *genitive of quantification* or GEN-Q for short (Babby 1987; Franks 1994; i.a.). Due to the numerous case and agreement options, Slavic quantified NPs – as the brief overview of the Slovenian pattern will show – present a grammatical system of significant complexity.

In addition to the type of QMOD, the case on the NP is also sensitive to number in Slovenian. When a singular NP is modified by a quantity adjective it gets external case, but with all other QMODs it gets GEN-Q. The case pattern is more complex with nonsingular NPs: (i) the NP gets external case with quantity adjectives, (ii) the NP gets GEN-Q with true quantifiers, but only if the external case is nominative or accusative, and (iii) the NP always gets GEN-Q with quantity nouns. As quantifiers show a greater flexibility in number than numerals, I begin the overview of the different case patterns in quantified NPs by looking at quantifiers.

---

34 In addition, colloquial Slovenian has an indefinite article homophonous with the numeral *en* 'one.M.SG' that also always inflects and has no restrictions on the number of the head noun it modifies. The article can even co-occur with numerals: *en-**i** dv-**e** punc-**i*** 'INDF-F.DU two-F.DU girl(F)-DU'.

The case pattern observed with quantity adjectives is illustrated by the examples in (45)–(46); note that there are differences in interpretation based on number which will be discussed below. Most importantly, the case of the quantifier and the case of the rest of the NP co-vary with both quantifiers because the entire NP is marked with the external case.

(45) a. *ves/vsa krompir* 'all.M.SG.NOM potato(M).SG.NOM' (mass/parts of entity)
 b. *vs-i krompirj-i* 'all-M.PL.NOM potato(M)-PL.NOM' (count)
 c. *ob-a krompirj-a* 'both-M.DU.NOM potatoes(M)-DU.NOM' (count)

(46) a. *vsak krompir* 'each/every.M.SG.NOM potato(M).SG.NOM' (entities)
 b. *vsak-i (tri-je) krompirj-i* 'every-M.PL.NOM (three-M.PL.NOM) potato(M)-PL.NOM' (sets)
 c. *vsak-a (dv-a) krompirj-a* 'every-M.DU.NOM (two-M.DU.NOM) potatoes(M)-DU.NOM' (sets)

The singular 'all' in (45a) either quantifies over a mass ('potato' can be count or mass in Slovenian) or over parts of a singular count entity ('the whole potato'), while the plural 'all' in (45b) quantifies over members of a plurality. In contrast, 'both' is always dual and quantifies over the members of a duality; see (45c). Distributive quantifiers like *vsak* 'each/every' in (46), which cannot quantify over a mass or parts of an entity, result in a different number split: singular ones quantify over entities (singularities), while nonsingular ones quantify over sets of entities (pluralities) (see Schwarzschild 1996: 77–9 regarding this contrast in English). The readings in (46) are respectively: each potato in the relevant set of potatoes (cf. (46a)) and each set of potatoes in the relevant set of sets of potatoes (cf. (46b,c)).[35] These splits unsurprisingly exclude pluralia tantum nouns, which require a plural quantifier even with the "singular readings" (e.g. *vsak-a vrat-a* 'each-N.PL door(N)-PL' = 'each door'); plural quantifiers behave in this respect just like nonsingular *en* 'one' (see Section 3.1).

The case pattern that results from modification by an uninflected true quantifier like *nekaj* 'some' is illustrated in (47); the instrumental examples stand in for all the *lexical cases* (i.e. not nominative or accusative), which all behave the same way regarding the GEN-Q/external case split.

(47) a. *nekaj krompirj-a* 'some potato(M)-SG.GEN' (singular, GEN-Q)
 b. *z nekaj krompirj-a* 'with some potato(M)-SG.GEN' (singular, GEN-Q)
 c. *nekaj krompirj-ev* 'some potato(M)-PL.GEN' (plural, GEN-Q)
 d. *z nekaj krompirj-i* 'with some potato(M)-PL.INS' (plural, external)

---

35 The numeral in parentheses denotes the cardinality of the set and is not needed if the cardinality is salient in the discourse. Since the cardinality of the set is always two with the dual, the numeral is then only used for emphasis.

Although 'some' shows the same number split in meaning as 'all' (quantification over a mass/parts of an entity vs. members of a plurality), the case pattern is different: a singular NP always gets GEN-Q, as shown in (47a,b), but a plural NP gets GEN-Q if the external case is *structural* (i.e. nominative and accusative) and external case if the external case is lexical, as shown in (47c,d).

It is useful to compare here the Slovenian GEN-Q pattern with the two most well known ones: the Russian one and the Bosnian/Croatian/Serbian (B/C/S) one (see Franks 1994, 1995: Ch. 4–5; Bošković 2006, 2008). In Russian, the case of NPs modified by true quantifiers is external with lexical cases and GEN-Q with structural cases. In B/C/S, on the other hand, the case with this type of modification is always GEN-Q, regardless of the external case. If we consider only plural NPs, Slovenian seems to behave like Russian. However, the two patterns come apart when we examine the case on *prequantifiers*: NP modifiers that can precede the QMOD, such as demonstrative determiners in (48); *pet* 'five' is also a true quantifier (see discussion of numerals below).

(48) 'these five beautiful girls'

    a. *èti*   *pjat'* *krasiv-yx*   *devušek*   [Russian]
       this.PL.NOM five   beautiful-PL.GEN girl(F).PL.GEN

    b. *ov-ih*   *pet lep-ih*   *devojak-a*   [B/C/S]
       this-PL.GEN five beautiful-PL.GEN girl(F)-PL.GEN

    c. *t-eh*   *pet lep-ih*   *punc*   [Slovenian]
       this-PL.GEN five beautiful-PL.GEN   girl(F).PL.GEN
       (Russian and B/C/S data from Franks 1995: 100)

In Russian, the demonstrative gets the external (nominative) case. In contrast, the demonstrative gets GEN-Q in both B/C/S and Slovenian – although in Slovenian it does get the external case in lexical case contexts. The Slovenian "mixed" pattern is in fact like the one in Czech, Slovak (Franks 1995: 135–7) and Polish – although the latter also marginally allows the Russian pattern (Franks 1995: 131–5). Another way in which these four languages differ from Russian and B/C/S is in relation to clausal agreement with GEN-Q marked NPs, as I will discuss in Section 3.4.1.

Returning to the different types of quantifiers, optionally inflected true quantifiers like *več* 'more' yield the same case pattern as uninflected ones, which is shown in (49). In fact, whenever the NP gets GEN-Q these also do not inflect; they can inflect only if the NP gets external case; see (49d).

(49) a. *več krompirj-a* 'more potato(M)-SG.GEN'   (singular, GEN-Q)
     b. *z več krompirj-a* 'with more potato(M)-SG.GEN'   (singular, GEN-Q)
     c. *več krompirj-ev* 'more potato(M)-PL.GEN'   (plural, GEN-Q)
     d. *z več(-imi) krompirj-i* 'with more(-M.PL.INS) potato(M)-PL.INS'
                                              (plural, external)

The remaining quantifier type are quantity nouns, like *večina* 'majority(F)'. These do not agree with the NP they quantify over because they have their own number and gender. We see in (50) that these quantifiers assign GEN-Q regardless of the number of the NP.

(50) a. *večin-a krompirj-a* 'majority(F)-SG.NOM potato(M)-PL.GEN' (singular, GEN-Q)
    b. *z večin-o krompirj-a* 'with majority(F)-SG.INS potato(M)-PL.GEN'
                                                                    (singular, GEN-Q)
    c. *večin-a krompirj-ev* 'majority(F)-SG.NOM potato(M)-PL.GEN' (plural, GEN-Q)
    d. *z večin-o krompirj-ev* 'with majority(F)-SG.INS potato(M)-PL.GEN'
                                                                    (plural, GEN-Q)

Although the quantifier is singular in all four examples, this is not always the only option. In the case of 'majority', nonsingular forms are not readily available due to the quantifier's meaning, but one can construct contexts where it can be nonsingular, such as (51).

(51) *z   dve-ma   večin-ama*   (dual QMOD, plural NP, GEN-Q)
    with two-DU.INS majority(F)-DU.INS
    glas-ov
    vote(M)-PL.GEN
    'with two majorities of votes (e.g. in both houses of parliament)'

Recall from Section 2.3.3 that this is the same pattern we get with mass nouns made countable via measure words, as in (52a). Additionally, we also see this pattern with fractions, as in (52b).

(52) a. *s   tre-mi   flaš-ami*   (plural QMOD, singular NP)
        with three-PL.INS bottle(F)-PL.INS
        vin-a
        wine(N)-SG.GEN
        'with three bottles of wine'

    b. *z   dve-ma   tretin-ama   glas-ov*   (dual QMOD, plural NP)
        with two-DU.INS third(F)-DU.INS vote(M)-PL.GEN
        'with two thirds of the votes'

Before considering the analysis of the case alternations, let us briefly turn to numerals, which divide into the same four types: (i) quantity adjectives ('one' through 'four'), (ii) optionally inflected true quantifiers ('five' and most higher numerals), (iii) uninflected true quantifiers ('zero', 'one million', and decimal numbers), and (iv) quantity nouns (fractions, thousands, millions, billions, …). In the case of complex numerals, it is the last segment that determines the inflection and case pattern. With "teens" ('11' to '19') and "tens" ('20' to '99'), it is the *base* (that which is added

to or multiplied) that comes last, as shown in (53a), while with bases of 100 and above, it is the *addend* (that which is added) that comes last: see (53b) for inflected numerals and (53c) for optionally inflected ones.

(53) a. **11**: *en-ajst točk* 'one-<u>teen</u> point(F).PL.GEN' (plural, GEN-Q)
     12: *dva-najst točk* 'two-<u>teen</u> point(F).PL.GEN' (plural, GEN-Q)
     31: *ena-in-tri-deset točk* 'one-and-three-<u>ten</u> point(F).PL.GEN' (plural, GEN-Q)
     32: *dva-in-tri-deset točk* 'two-and-three-<u>ten</u> point(F).PL.GEN' (plural, GEN-Q)
   b. **101**: *sto en-a točk-a* 'hundred <u>one</u>-F.SG point(F)-SG' (singular, external)
     102: *sto dv-e točk-i* 'hundred <u>two</u>-F.DU point(F)-DU' (dual, external)
     103: *sto tri točk-e* 'hundred <u>three</u>.F.PL point(F)-PL' (plural, external)
   c. **105**: *sto pet točk* 'hundred <u>five</u> point(F)-PL.GEN' (plural, GEN-Q)
     110: *sto deset točk* 'hundred <u>ten</u> point(F)-PL.GEN' (plural, GEN-Q)

Excluding quantity adjectives, which we already saw in Section 3.1, the case patterns for the different types of numerals are illustrated in (54) (lexical cases are again illustrated with instrumental case).

(54) a. **5&up, thousand₁:** (*true quantifiers – optionally inflected*)
     – structural: *pet točk* 'five point(F).PL.GEN', *tisoč točk* 'thousand point(F).PL.GEN', ...
     – lexical: *s pet(-imi) točk-ami* 'with five(-PL.INS) point(F).PL.INS', *s tisoč(-imi) točk-ami* 'with thousand(-PL.INS) point(F).PL.INS', ...
   b. **zero, million₁, decimal numbers:** (*true quantifiers – uninflected*)
     – structural: *nič točk* 'zero point(F)PL.GEN', *miljon točk* 'million point(F)PL.GEN', ...
     – lexical: *z nič točk-ami* 'with zero point(F)PL.INS', *z miljon točk-ami* 'with million point(F)-PL.INS', ...
   c. **thousand(s)₂, million(s)₂, billion(s), ...:** (*quantity nouns*)
     – structural(PL): *tisoč-i točk* 'thousand(M)-PL.NOM point(F).PL.GEN', *miljon-i točk* 'million(M)-PL.NOM point(F)PL.GEN', ...
     – lexical(SG): *s tisoč-em točk* 'with thousand(M)-SG.INS point(F).PL.GEN', *z miljon-om točk* 'with million(M)-SG.INS point(F)PL.GEN', ...
     – lexical(PL): *s tisoč-i točk* 'with thousand(M)-PL.INS point(F).PL.GEN', *z miljon-i točk* 'with million(M)-PL.INS point(F)PL.GEN', ...

Note that 'thousand' and 'million' have two variants, where the quantity noun variant can have different number values just like quantity noun quantifiers (singular structural case forms are not given in (54c) as they are indistinguishable from their optionally inflected/uninflected counterparts).[36]

---

[36] In complex numerals, only one of the two variants is used as the base, although interestingly not the same one: cf. *pet <u>tisoč</u>* 'five thousand' (optionally inflected) vs. *pet <u>miljon-ov</u>* 'five million(M)-PL.GEN (nominal).

The one difference between quantifier and numeral modification is the lack of mass and parts of entity readings (in the absence of measure words), which is expected given that other than 'one' all numerals must quantify over a plurality or duality. As for the singular versus nonsingular semantic split exemplified by *vsak* 'each', where singular forms get a quantification over entities reading, while plural and dual forms get a quantification over sets of entities reading, we saw already in Section 3.1 that the same split arises with the numeral *en* 'one', which can be singular or nonsingular.

There are many analyses of Slavic quantified NPs in the generative tradition, most of them building on the insights of Babby's (1987) analysis of Russian; namely, that the structure of quantified NPs is different when they receive GEN-Q as opposed to external case. There are, however, also several key differences in implementation between specific analyses, such as for example Franks (1994, 1995) and Bošković (2006, 2008). For reasons of space, I only briefly outline these two approaches and where they differ, also in relation to their applicability to Slovenian.

Bošković (2006, 2008) argues that GEN-Q is assigned in structures like (55a), where the NP is the complement of the functional syntactic head (F for convenience sake) that hosts only true quantifiers (henceforth Qs) in its specifier, while the NP gets external case in the structure in (55b), where QMODs are quantity adjectives (henceforth A_Qs) in the specifier of NP.

(55) a. [V … [$_{FP}$ Q [$_{F'}$ F [$_{NP}$ N]$_{GEN-Q}$]]]

   b. [V … [$_{FP}$ F [$_{NP}$ A_Q [$_{N'}$ N]]]]

As shown in (55a), F is also the head that assigns GEN-Q, but crucially only when external case is not assigned to the NP. Franks' (1994, 1995) analysis does not differ much here, the main difference is that GEN-Q is assigned directly by Q.[37] Crucially, both GEN-Q and external case are assigned to the whole NP and expressed on all inflected elements inside the NP.

The issue lies in determining when each structure is used based on the QMOD and the external case, but also how this can differ across languages. In the case of B/C/S, where GEN-Q is never overridden by the external case, this is rather straightforward: Bošković (2008) suggests that Qs always take the GEN-Q structure in (55a) and A_Qs always take the structure in (55b). But this cannot work for Slovenian and Russian, where (55b) must be used even in GEN-Q contexts if the external case is lexical. The consensus is that this follows from lexical cases being *inherent cases*

---

37 Although, Franks (1994, 1995) does assume a head in the same position as F does assign GEN-Q only in the case of distributive *po*-constructions in Russian; see Franks (1994: 633–55, 1995: 139–157) for the details.

(Chomsky 1986), and thus either their assignment has precedence over structural case assignment (Franks 1994, 1995) or they never compete with structural cases due to their role in licensing thematic relations (Bošković 2006, 2008). The key idea is that GEN-Q is a structural case and can thus be overridden only if the external case is inherent. Although, Franks (1994, 1995) argues that inherent GEN-Q also exists. In fact, he proposes that GEN-Q is inherent in B/C/S and that this is why it cannot be overridden (more on this below). In contrast, for Bošković (2006, 2008) overriding GEN-Q means picking the (55b) structure rather than (55a) even if the QMOD stays the same. This means that GEN-Q assigning QMODs in Russian and Slovenian must come in both Q and A_Q versions.

There is actually independent evidence for the two structures in Slovenian: clitic pronouns do not take specifiers, but they can receive GEN-Q (Golden and Sheppard 2008, 2009), as in (56).

(56) *Včeraj* **jih**$_i$ *je* [*pet* t_i] *prodal* *zjutraj.*
yesterday 3PL.GEN AUX.3SG five sold.M.SG in.morning
'Yesterday he sold five of them in the morning.'

Since 'five' cannot be a specifier here, the pronoun must be a complement in the quantified phrase before it moves to its surface position (due to being a second position clitic; see Section 2.2.3).

We expect then that with externally case marked quantified NPs, the NP should not be replaceable with a clitic pronoun, which is borne out, as the examples in (57) show.[38, 39] Note that omitting the clitic is grammatical with a noun ellipsis parse (cf. 'I sold one car and he sold two'), and that noun ellipsis is possible with the optionally inflected numeral in (57b) only if it is inflected.[40]

---

38 B/C/S shows a similar split in pronominalization (Giusti and Leko 1995, 2005), although without the (56) vs. (57b) contrast, since GEN-Q is never overridden. Russian and Polish show the same split as Slovenian, although in relation to NP-internal pronoun movement (Giusti and Leko 2005: 157, 169–70). Czech behaves like Slovenian, except that the clitic pronoun is optional in the counterparts of (56) and (57b) (Veselovská 1995: 236–7(n23)); however, Giusti and Leko (2005: 166) tentatively analyze these as involving secondary predication rather than modification.

39 Speakers find the pronouns strongly degraded but not fully ungrammatical, which is crucially also the case with uncontroversially adjectival modifiers. Some singular quantifiers can co-occur with a clitic pronoun (e.g. *cel* 'whole', *ves* 'all'); these must be accusative with masculine pronouns even in non-accusative contexts, which is also the case with secondary predicates of clitic pronouns (Perlmutter and Orešnik 1973: 439–41) (cf. Czech in footnote 38).

40 This can be related to the proposal of Lobeck (1990, 1995) and Saito and Murasugi (1990a,b) that complements of functional heads can be elided only if the functional head agrees with its specifier. This would imply that the structure of NPs is more complex: A_Qs must be specifiers of a functional head below FP in the extended projection of the NP.

(57) a. *Včeraj* (?****ju**$_i$*) *je* [*dv-**a*** ___$_i$] *prodal*
    yesterday   3DU.ACC   AUX.3SG   two-M.DU.ACC   sold.M.SG
    zjutraj.
    in.morning
    'Yesterday he sold two (of them) in the morning.'

   b. *Včeraj* (?****jim**$_i$*) *je* [*pet**(-**im***) ___$_i$] *pomagal*
    yesterday   3PL.DAT   AUX.3SG   five-PL.DAT   helped.M.SG
    zjutraj.
    in.morning
    'Yesterday he helped five (of them) in the morning.'

Importantly, not only is the pronoun disallowed with A$_Q$s, as in (57a), it is also disallowed with otherwise GEN-Q inducing Qs if the external case is lexical, as in (57b), where the verb *pomagati* 'help' assigns dative case to the object. This fits the idea that the same QMOD alternates between Q and A$_Q$ forms and therefore the two NP structures in (55) depending on the external case.

    Although a Russian-style analysis appears to be extendable to Slovenian, recall that the case on prequantifiers is different in the two languages. In Russian, these bear the external case in GEN-Q contexts, while in Slovenian and B/C/S they bear GEN-Q. Both Franks and Bošković here build on Corbett's (1979) proposal that the prequantifier is introduced below the GEN-Q assigner and moves to its surface position, as illustrated in (58) with nominative as the external case.

(58) [Dem$_{GEN-Q/NOM}$ [$_{FP}$ Q [$_{F'}$ F [$_{NP}$ *t* [N]]$_{GEN-Q}$]]]

The idea is that prequantifiers can be assigned GEN-Q as part of the NP before they are displaced, which derives the B/C/S and Slovenian pattern. As for Russian, Franks (1994, 1995) argues that GEN-Q is assigned only after the prequantifier moves out of the NP, because GEN-Q is a structural case in Russian, unlike in B/C/S, where it is inherent. Bošković (2008), on the other hand, suggests that case bearing elements may receive more than one case in Russian (more on this below): the prequantifier receives GEN-Q in situ and external case in its derived position, but only the latter is realized for morphological reasons. Crucially, all three languages also allow prequantifiers to surface in their in situ position, in which case even Russian ones bear GEN-Q.

    Slovenian is relevant here because it seems to present a paradox in Franks' analysis: GEN-Q must be structural since it can be overridden, and GEN-Q must be inherent since it is assigned to prequantifiers prior to movement. In order to resolve this, Franks (1995: 132–3, 137) proposes that in Slovenian, as well as Polish, Czech, and Slovak, which all have the same pattern, GEN-Q is inherent, but the reason why it can be overridden is different than in Russian: GEN-Q can be assigned to an NP

only if the NP is accusative. When the external case is inherent, GEN-Q cannot be assigned, since inherent case assignment has precedence over structural accusative case assignment. The caveat is that in nominative assigning contexts the NP must be assigned accusative case by stipulation. There is, however, some potential evidence for this in Slovenian, which I discuss in Section 3.4.1.

In contrast, Bošković (2008) attributes the lack of GEN-Q on prequantifiers in Russian to multiple case assignment. In Russian passives (Franks 1995: 349), but not B/C/S passives (Bošković 2008: 285), lexical cases can convert to nominative case under promotion to subject. Bošković takes this to mean that the NP gets lexical case in situ before it moves to subject position and also gets nominative case, which is the case that gets morphologically realized. The idea is that a general availability of multiple case assignment is behind both the prequantifier and the passive case facts. This analysis is also compatible with Slovenian, given that it has both GEN-Q prequantifiers and disallows lexical-to-nominative case conversion in passives: structural accusative case converts to nominative in passives, as shown in (59a),[41] but the lexical dative case assigned to the theme/patient argument by the verb *pomagati* 'help' remains dative also in passives, as shown in (59b).

(59) a. *Pit-e   so   se   pojedle.*
        pie(F)-PL.NOM AUX.3PL REFL eaten.F.PL
        'The pies were eaten.'

    b. *An-i   se   je   pomagalo.*
        Ana-SG.DAT REFL AUX.3SG helped.N.SG
        'Ana was helped.'

Franks (1994, 1995) and Bošković (2006, 2008) consider in detail only numerals, where the NP's number is fixed. However, recall that in Slovenian some quantifiers modify plural and singular NPs, and that if a Q modifies a singular NP its GEN-Q cannot be overridden. In such cases, demonstratives also bear GEN-Q even in lexical case contexts, whether they precede or follow the Q:

(60) *s/z*   { *t-ega*   } *malo* { *t-ega*   } *krompir-ja*
     with   this-M.SG.GEN   little   this-M.SG.GEN   potato(M)-SG.GEN
     'with this little bit of potato/with a little bit of this potato'

Quantity nouns (henceforth N_Qs), whose number is independent from the NP, are the other context where GEN-Q cannot be overridden. In such cases, either the N_Q,

---

41 The examples in (59) are impersonal passives. Slovenian additionally has personal participal passives (see (74), (75) and (77) in Section 3.4.1), which also allow accusative-to-nominative but not lexical-to-nominative case conversion; in fact, monotransitive sentences with a lexically case marked object have no personal passive counterparts.

the quantified NP, or both can be modified by a demonstrative, each bearing the case of the element it modifies, as shown in (61).

(61) *s*   (*t-o*   ) *četrtin-o*   (*t-ega*   ) *krompir-ja*
      with   this-F.SG.INS   fourth(F)-SG.INS   this-M.SG.GEN   potato(M)-SG.GEN
      'with a/this fourth of a/this potato'

The clitic pronoun test used above is also helpful here to further probe the structure of these two QMOD types. Although the pronoun is always genitive with N_Qs, it is allowed with structural but not with lexical external cases regardless of its number, as shown in (62).[42]

(62) a. *Janez **jih**$_i$*   / ***je**$_i$*   *je*   [*večin-o* t_i] *dal*   *v*
        Janez 3PL.GEN / 3SG.F.GEN AUX.3SG   most(F)-SG.ACC   put.M.SG in
        omak-o.
        sauce(F)-SG.ACC
        'Janez put most of them/it to the sauce.'

      b. **Janez **jih**$_i$*   / ***je**$_i$*   *je*   [*večin-i* t_i] *dodal*
         Janez 3PL.GEN / 3SG.F.GEN AUX.3SG   most(F)-SG.DAT   added.M.SG
         žvepl-o.
         sulfur(F)-SG.ACC
         'Janez added sulfur to most of them/it.'

We observe the same pattern with Qs modifying singular NPs, as shown in (63).

(63) a. *Janez **je**$_i$*   *je*   [*nekaj* t_i]$_{ACC}$ *dal*   *v omak-o.*
        Janez 3SG.F.GEN AUX.3SG   some   put.M.SG in sauce(F)-SG.ACC
        'Janez put some of it to the sauce.'

      b. **Janez **je**$_i$*   *je*   [*nekaj* t_i]$_{DAT}$ *dodal*   *žvepl-o.*
         Janez 3SG.F.GEN AUX.3SG   some   added.M.SG sulfur(N)-SG.ACC
         'Janez added sulfur to some of it.'

However, the reason why the clitic pronoun is banned in (63b) might be different than in (62b). As shown in (64), Qs with a GEN-Q complement are – unlike N_Qs – independently banned in lexical case contexts, unless the case is assigned by a preposition (Golden and Sheppard 2008, 2009).

---

42 Although, consistent with the findings of Giusti and Leko (1995, 2005), noun ellipsis is possible in both contexts.

(64) *Janez je* [*večin-**i*** / **nekaj mešanic-**e***] *dodal*
Janez AUX.3SG   most(F)-SG.DAT / some mixture(F)-SG.GEN   added.M.SG
žvepl-o.
sulfur(N)-SG.ACC
'Janez added sulfur to most/some of the mixture.'

The same restriction is observed in B/C/S, although not only with singular NPs, since GEN-Q cannot be overridden regardless of the NP's number. Bošković (2008) takes this as evidence that inherent (lexical) case assignment is a prerequisite for licensing the thematic role it is associated with: no part of the NP bears the inherent case, so the NP cannot be theta-marked.

I suggest that with N_Qs and related cases like measure words, the genitive case on the NP is not GEN-Q but the genitive normally assigned to noun complements in Slovenian also outside quantified NPs.[43] As shown in (65), the N_Q takes the quantified NP as a complement, and either NP can be modified by its own demonstrative marked with the same case as the NP it modifies.

(65) [$_{NP_1}$ (Dem$_1$) N$_Q$ [$_{NP_2}$ (Dem$_2$) N]$_{GEN}$]   (quantity noun construction)

Since the quantified NP is always a complement, it can be substituted for a clitic pronoun if the external case is structural. The impossibility of this with lexical external cases can be attributed to the ban on extraction out of inherently case marked elements (Starke 2001: 38–40; Bošković 2018): the clitic pronoun must extract to its surface position outside the NP, but cannot do so because the NP is inherently case marked. This issue crucially does not arise in cases where quantified clitic pronouns are disallowed to begin with, such as A_Q constructions.

The situation with Qs and singular NPs is less straightforward, as there are in principle two analyses compatible with the facts. One option is that only with singular NPs, the GEN-Q inducing QMODs lack an A_Q counterpart just like in B/C/S, so the NP structure is always the one in (66a). The other option is that the NP is the complement of a covert N_Q modified by the Q, as shown in (66b), where the NP gets (non-GEN-Q) genitive case the same way as in (65).

(66) a. [$_{FP}$ Q$_{SG}$ [$_{F'}$ F [$_{NP}$ (Dem) N$_{SG}$]$_{GEN-Q}$]]
      b. [$_{FP}$ Q [$_{F'}$ F [$_{NP_1}$ ⟨N$_Q$⟩ [$_{NP_2}$ (Dem) N$_{SG}$]$_{GEN}$]]]

The idea behind (66b) is that the covert N_Q is tied to the interpretation of the singular NP. Recall that the relevant interpretations here are quantification over a mass and quantification over parts of the entity denoted by the NP, so the covert N_Q essentially functions as a measure word. Possible evidence for this analysis comes from the fact that overt N_Qs can also be modified by Qs: *nekaj desetin pit-**e*** 'some

---

43 See also Leko (1986: 170–81, 1989: 34–42) regarding this type of genitive case assignment in B/C/S.

tenth(F).PL.GEN pie(F)-SG.GEN' ('a few tenths of the pie'). However, since A_Qs can also modify overt N_Qs, one should find a principled reason to exclude covert N_Qs in those cases. The restriction could be tied to the precise meaning of the covert N_Q – recall that some A_Q quantifiers already show a different semantic split based on the number of the NP.

A similar issue arises with the analysis in (66b), where the unavailability of an A_Q counterpart of Q must be tied to singular NPs. This could be implemented in terms of number agreement between the Q and the NP, although at least with uninflected Qs there is no direct evidence for this.

What we can conclude is that the Slovenian data largely fits the standard analyses of case in Slavic quantified NPs. At the same time, there remain open questions when it comes to the difference in the status of GEN-Q between plural and singular NP. This topic has so far not received much attention, partly because of the focus in the literature on numeral modification and partly because the relevant contrast is not observed in all the other Slavic languages.

3.3 Verbal inflection

Before turning to the role of number in clausal agreement, let us briefly take a closer look at verbal inflectional morphology. Main verbs as well as auxiliary verbs show a three-way number distinction. This is illustrated respectively by the present tense main verbs in (67) and the auxiliary verbs functioning as copulas in the predicative adjective constructions in (68).

(67) a. *Pride-m*   (*jaz*) *tja.*
       come-1SG (I)   to.there
       'I'll come there.'

   b. *Pride-va*   (*midva*) *tja.*
       come-1DU (we.DU) to.there
       'We'll (both) come there.'

   c. *Pride-mo* (*mi*)   *tja.*
       come-1PL (we.PL)   to.there
       'We'll come there.'

(68) a. *On **je***   *star.*
       he AUX.3SG old.M.SG
       'He's old.'

   b. *Onadva*   ***sta***   *star-**a**.*
       they(M).DU AUX.3DU old-M.DU
       'The two of them are old.'

c. *Oni* ***so*** *star-**i***.
   they(M).PL AUX.3PL old-M.PL
   'They are old.'

Slovenian is also a null subject language, so agreement is often the only expression of number in a clause (cf. (67)); overt subjects are only used when discourse-prominent or emphatic. As seen in (68), the adjectives used predicatively with an auxiliary also agree in number with the subject; this includes other elements with adjectival inflection, like the numeral in (69a) and possessor in (69b).

(69) a. *Pot-i* ***sta*** *dv-e.*
       paths(F)-DU AUX.3DU two-F.DU
       'There are two paths / The paths are two.'

   b. *Knjig-i* ***sta*** *moj-i.*
       book(F)-DU AUX.3DU my-F.DU
       'The (two) books are mine.'

The verbal inflection patterns are presented in Table 20; notice that 2nd and 3rd person are neutralized throughout all person marking forms in the dual (compatible with the markedness characterization from Section 2.2.2). Predicative adjectives in-

Tab. 20: Verbal agreement.

		SG	**PL**	**DU**
present tense *delati* ('to work')	1	del**am**	dela**mo**	dela**va**
	2	dela**š**	dela**te**	dela**ta**
	3	dela	dela**jo**	
present tense AUXILIARY *biti* ('be')	1	se**m**	s**mo**	s**va**
	2	si	s**te**	s**ta**
	3	je	s**o**	
future tense AUXILIARY *biti* ('be')	1	bo**m**	bo**mo**	bo**va**
	2	bo**š**	bo**ste**	bo**sta**
	3	bo	bo**jo**	
imperative *delati* ('to work')	1		delaj**mo**	delaj**va**
	2	delaj	delaj**te**	delaj**ta**
	3			
participle *delati* ('to work')	M	delal	delal**i**	delal**a**
	F	delal**a**	delal**e**	delal**i**
	N	delal**o**	delal**a**	

flect the same as attributive adjectives, apart from lacking specificity marking in masculine singular forms (i.e. the **-i** suffix; see Section 3.1).

Having presented the inflectional suffix forms for all agreeing elements, they will be segmented out in the examples below in order to facilitate comparison between different forms.

3.4 Constraints on agreement

The main factor constraining clausal agreement in Slovenian is the case of the agreement controller (the element which determines the agreement), although the syntactic category of the agreement target (the element whose form is determined by the agreement) also plays a role. Additionally, even though word order does not generally play a role in agreement, it can impose certain restrictions when agreement takes place with conjoined elements.

3.4.1 Agreement with quantified noun phrases

The different case patterns on quantified NPs discussed in Section 3.1 crucially also affect agreement with those NPs. For example, with quantity adjectives (A_Qs) like *dva* 'two' and *oba* 'both' or *tri* 'three' and *ves/vsa* 'all' we see a regular clausal agreement pattern, as illustrated in (70).

(70) a. [*Dv-**e*** / *Ob-**e***   pisateljic-***i***] s-***ta***   prišl-***i***.   (*dual*)
       two-F.DU / both-F.DU writer(F)-DU   AUX-3DU arrived-F.DU
       'Two/Both writers(F) have arrived.'

     b. [*Tri*   / *Vs-**e***   pisateljic-***e***] s-***o***   prišl-***e***.   (*plural*)
       three.F.PL / all-F.PL writer(F)-PL   AUX-3PL arrived-F.PL
       'Three/All the writers(F) have arrived.'

In contrast, with true quantifiers (Qs), like *pet* 'five' and *več* 'more', which yield GEN-Q case on the NP, we see default neuter singular agreement on verbs, as in (71).

(71) [*Pet* / *Več*   pisateljic   ] ***je***   prišl-***o***.   (*default singular*)
       five / more writer(F).PL.GEN   AUX.3SG arrived-N.SG
       'Five/More writers(F) have arrived.'

Finally, with quantity nouns (N_Qs), like *petin-**a*** 'fifth(F)-SG' and *večin-**a*** 'majority(F)-SG', the agreement is not with the quantified NP, but with the N_Q, as shown in (72), where we can see that the main verb agrees with the numeral/quantifier in both number and gender.

(72) [*Petin-a* / *Večin-a* *pisatelj-ev*] ***je*** (*singular*)
    fifth(F)-SG / majority(F)-SG writer(M)-PL.GEN AUX.3SG
prišl-a.
arrived-F.SG
'A fifth/The majority of the writers(M) have arrived.'

Note that this pattern is consistent with the N$_Q$ being the head of the projection that contains the quantified NP, as was suggested previously in Section 3.1.

A third case pattern for quantified NPs appears to be possible in very specific contexts. Namely, if a subject NP modified by a numeral is a unit of time, distance, money, or any other conventionalized unit of measurement, the NP can be either nominative or accusative, as illustrated in (73) with *teden* 'week(M).SG' as the unit of time. In such cases, the nominative option yields regular agreement, as in (73a), and the accusative option yields default singular agreement, as in (73b).

(73) a. [*Tri-je* *tedn-i*] *s-o* *šl-i* *v*
    three-M.PL.NOM week(M)-PL.NOM AUX-3PL gone-M.PL in
nič.
nothing(M).SG.ACC

   b. [*Tri* *tedn-e*] ***je*** *šl-o* *v*
    three.M.PL.ACC week(M)-PL.ACC AUX.3SG gone-N.SG in
nič.
nothing(M).SG.ACC
'Three weeks have gone to waste.'

The patterns exemplified in (70)–(73) fit a single generalization: agreement occurs with a quantified NP if and only if the head of the phrase is nominative. However, we will see that the generalization is not quite right if predicative adjectives (and related modifiers) are also considered.

In complex numerals, where the last numeral is the head (given that it determines the inflection and case pattern for the modified NP), the clausal agreement pattern is also determined this way. Thus, as shown in (74), '101' yields singular, '102' dual, and '103' plural agreement.

(74) a. *Prodan-a **je*** *bil-a* [*sto* *en-a* *knjig-a*]. (*singular*)
    sold-F.SG AUX.3SG been-F.SG hundred one-F.SG book(F)-SG
'One hundred and one books were sold.'

   b. *Prodan-i* *s-ta* *bil-i* [*sto* *dv-e* *knjig-i*]. (*dual*)
    sold-F.DU AUX-3DU been-F.DU hundred two-F.DU book(F)-DU
'One hundred and two books were sold.'

c. *Prodan-e s-o bil-e* [*sto tri knjig-e*]. (*plural*)
   sold-F.PL AUX-3PL been-F.PL hundred three.F.PL book(F)-PL
   'One hundred and three books were sold.'

Note that the present participle, which is the main verb in participial passives like those in (74), also agrees with the subject in number. Crucially, the present participle is morphologically an adjective and in addition to number and gender also inflects for case. We see this clearly when the subject NP is modified by a Q numeral. A complex numeral like '105' yields default verbal agreement because its head is a Q (cf. (71)), but note that in (75) '105' only yields default agreement on the auxiliaries and not on the participle, which agrees with the head noun in number and case.

(75) *Prodan-ih je bil-o* [*sto pet* (*default + plural*)
    sold-PL.GEN AUX.3SG been-N.SG hundred five
    knjig].
    book(F).PL.GEN
    'One hundred and five books were sold.'

The same agreement is also observed with other predicative adjectives, as seen in (76). In addition to primary predicate cases like (76a), this also includes secondary predication like in (76b); note that the adjective agrees with any argument of which it is predicated, in this case a direct object.

(76) a. [*Pet knjig*] *je rasprodan-ih.* (*default + plural*)
      five book(F).PL.GEN AUX.3SG sold.out-F.PL.GEN
      'Five books are sold out.'

    b. [*Pet knjig*] *si vrnil* (*plural-object*)
      five book(F).PL.GEN AUX.2SG returned-M.SG
      raztrgan-ih.
      torn-F.PL.GEN
      'You returned five books torn.'

Thus, the generalization concerning agreement with quantified NPs should be revised to: (i) verbs agree with a quantified NP if and only if the head of the phrase is nominative, and (ii) other predicates agree with a quantified NP if and only if they are predicated of it.

    The half-default-half-full agreement observed in cases like (75) and (76a) is also found in Polish (Dziwirek 1990), but crucially not in Russian and B/C/S, where GEN-Q subjects yield either default or plural clausal agreement, but the agreement on verbs and other predicates is the same in both cases (Franks 1994: 613–26, 1995: 106–17).

What sets the Slovenian pattern apart is that verbs never agree with GEN-Q NPs but other types of predicates always agree with them.

Incidentally, case agreement with non-verbal predicates also reveals a complication concerning the apparent third quantified NP case pattern, where the NP can be either nominative or accusative. With the accusative option we would expect the present participle to be accusative, but this is not what we get. As shown in (77), the participle must be genitive in such cases.

(77) *Ukraden-**ih**/*-**e***   *je*   *bil-**o***   [*sto*   *tri*
     stolen-PL.GEN/-M.PL.ACC AUX.3SG been-N.SG   hundred three.M.SG.ACC
     *evr-**e***   ].
     Euro(M)-PL.ACC
     'One hundred and three Euros were stolen.'

One way to interpret (77) is that 'Euros' is not the head of the NP, the real head is a covert/elided element: an abstract plural noun whose quantity is being measured (e.g. 'three Euros of MONEYS') and which is assigned GEN-Q. This would mean the nominative/accusative alternation (cf. (73)) is really an alternation between two NP structures: one where the unit word, like 'Euro' or 'week', is the head (nominative NP) and one where the abstract noun is the head (accusative NP). In the latter, the numeral and the unit word together form a complex quantifier that assigns GEN-Q.

If this interpretation is right, it provides indirect evidence for Franks' (1995) analysis of GEN-Q in Polish, Czech, Slovak, and Slovenian. As noted in Section 3.2, Franks argues that in these languages GEN-Q is only assigned in accusative NPs and that even otherwise nominative subjects are accusative when the NP gets GEN-Q. This is also argued to be the reason why verbs cannot agree with such NPs, as verbs only agree with nominative NPs. There is some direct evidence for this claim in Polish, where Q numerals like 'five' bear accusative morphology even in nominative contexts.[44] However, in Slovenian (as well as Czech and Slovak; see Franks 1995: 137), where the corresponding Q numerals are uninflected in the relevant contexts, it is impossible to tell what case they bear. Now consider (77) with the abstract noun analysis in mind: the modifier is accusative and the head noun gets GEN-Q, which is exactly the case pattern predicted by Franks.

---

44 Technically, the numeral has a nominative form (*pięć*) with feminine nouns and a genitive form (*pięciu*) with masculine human nouns, but this are exactly the forms one expects in an accusative noun context, since Polish shows NOM-ACC syncretism in the former context and ACC-GEN syncretism in the latter (Franks 1995: 132).

3.4.2 Agreement and coordination

In the examples discussed in the previous section, the primary factor constraining agreement was case, not word order. The agreement controller in those examples either precedes or follows the agreement target, due to the relatively free word order of major syntactic constituents in Slovenian, but that has no effect on agreement. The position of agreeing elements in relation to each other does, however, play a role when agreement occurs with coordinated words and phrases. The basic principle behind this can be observed, at least superficially, when coordinated numerals modify a single noun and only the numeral closest to the noun agrees with it:

(78) *en-a ali dv-e miš-i* 'one-F.SG or two-F.DU mouse(F)-DU'
      dv-e ali en-a miš 'two-F.DU or one-F.SG mouse(F).SG'

But the pattern in (78) is in principle compatible with an ellipsis analysis where there is no two-to-one relation between numerals and the noun: *en-a (miš) ali dv-e miš-i* 'one-F.SG (mouse(F).SG) or two-F.DU mouse(F)-DU'. To avoid this confound, we must look at so-called *right node raising* contexts like (79), which do not involve ellipsis (see Shen 2018: 81–9).

(79) a. *T-a   in tist-a   punc-a   <u>s-ta</u>   par.*
         this-F.SG and that-F.SG girl(F)-SG AUX-3DU couple(M).SG

      b. **T-a   in tist-a   punc-i   <u>s-ta</u>   par.*
         this-F.SG and that-F.SG girl(F)-DU AUX-3DU couple(M).SG
         'This girl and that girl are a couple.'
         (Shen 2018: 220)

The demonstratives are each individually in agreement with the singular noun. This is in sharp contrast with the verb, which shows *summative agreement*: agreement with the conjunction as a whole. In Slovenian and other languages with a dual, summative agreement with two singular conjuncts is dual, while all other number combinations of conjuncts yield plural summative agreement (Corbett 2000: 198). Different agreeing elements thus behave differently when it comes to agreement with conjunctions; see Shen (2018) for a more general discussion of such discrepancies.

The relevance of linear order for agreement is much clearer with the so-called *Closest Conjunct Agreement* (CCA) effect (Koutsoudas 1968; Corbett 1983: Ch. 7; Nevins and Weisser 2019; i.a.), where clausal agreement occurs only with the conjunct closest to a predicate. Slovenian data have been instrumental in the study of this phenomenon (Corbett 1983: Ch. 7; Marušič, Nevins, and Saksida 2007; Marušič, Nevins, and Badecker 2015; Arsenijević et al. 2020; i.a.), although primarily in relation to gender agreement. For example, if two conjuncts are plural and differ in gender, as in (80a), agreement can *resolve* to masculine plural (note that neither conjunct is

masculine in the example), or it can occur with either conjunct;[45] agreement with the last conjunct is here a case of CCA. If instead the two conjuncts are singular, as in (80b), number agreement is summative and gender agreement resolves to masculine.

(80) a. [*Krav-e   in   telet-a*   ] *s-o   se   pasl-i/-a/-e.*
       cow(F)-PL and calf(N)-PL   AUX-3PL REFL graze-M.PL/-N.PL/-F.PL
       'Cows and calves grazed.'
       (Marušič and Nevins 2010: 303)

   b. [*Fant   in   dekl-e*   ] *s-ta   brcal-a   žog-o.*
      boy(M).SG and girl(N)-SG   AUX-3DU kicked-M.DU ball(F)-SG.ACC
      'A boy and a girl were kicking a ball.'
      (Marušič and Nevins 2010: 309)

Crucially, CCA can be observed for number agreement in Slovenian if one of the conjuncts is a NP with GEN-Q (Marušič and Nevins 2010). If the other conjunct is closest to the verb, as in (81a), then the verb agrees with that conjunct in number and gender. If the GEN-Q conjunct is closest instead, either because the conjunction follows the verb, as in (81b), or because the conjuncts are reversed, as in (81c), the verb shows default neuter singular agreement.

(81) a. [*Šest fant-ov   in   štir-i   punc-e*   ] *s-o   plesal-e.*
       six   boy(M)-PL.GEN and four-F.PL girl(F)-PL   AUX-3PL danced-F.PL

   b. *Plesal-o   je   [šest fant-ov   in   štir-i   punc-e   ].*
      danced-N.SG AUX.3SG   six boy(M)-PL.GEN and four-F.PL girl(F)-PL
      'Six boys and four girls danced.'

   c. [*Štir-i   punc-e   in   šest fant-ov*   ] *je   plesal-o.*
      four-F.PL girl(F)-PL and six   boy(M)-PL.GEN   AUX.3SG danced-N.SG
      'Four girls and six boys danced.'

Experimental data presented in Marušič and Nevins (2010) and Marušič, Nevins, and Badecker (2015) show that CCA is strongly preferred in such contexts, but it is not the only option. In fact, if interpretation is controlled for (which was impossible in the experimental studies in question), a resolution pattern is more easily attainable. For example, if a reciprocal interpretation is forced with 'each other', a resolution with masculine plural agreement becomes the only option in (82), and an acceptable alternative to default agreement when the GEN-Q conjunct is the one closest to the verb.

---

45 With uniform-gender conjuncts the verb normally agrees, as expected, for the shared gender, although masculine plural gender resolution is not entirely ruled out in such cases, as observed by Marušič, Nevins, and Badecker (2015).

(82) [*Šest fant-ov*   *in*   *štir-i*   *punc-e*   ] ...
     six   boy(M)-PL.GEN and four-F.PL girl(F)-PL

     a. ... *s-o*   *plesal-i*   *drug*   *z*   *drug-im.*   (*resolution*)
             AUX-3PL danced-M.PL each.M.SG with other-M.SG.INS

     b. ... **s-o*   *plesal-e*   *drug(-a)*   *z*   *drug-im/-o.*   (*CCA*)
             AUX-3PL danced-F.PL each.M.SG(-F.SG) with other-M.SG.INS/-F.SG.INS
             'Six boys and four girls danced with each other.'

This is reminiscent of a pattern Franks (1994: 660–1) discusses for Russian agreement. Recall that in Russian agreement is typically either default or plural if the subject is a NP with GEN-Q. However, when 'each other' is used, agreement must be plural (i.e. "semantic" agreement). What is interesting is that in Slovenian this effect is limited to conjoined subjects like the one in (82).

     In cases with one singular and one GEN-Q conjunct we find yet another pattern. If the singular conjunct is closest, we see either masculine plural agreement or default agreement:

(83) [*Peg fant-ov*   *in*   *en-a*   *punc-a*   ] *s-o*   *plesal-i*   /
     five boy(M)-PL.GEN and one-F.SG girl(F)-SG   AUX-3PL danced-M.PL /
     je   *plesal-o.*
     AUX.3SG danced-N.SG
     'Five boys and one girl danced.'

Neither option here is CCA: the first is summative number agreement with gender resolution and the second is default either in relation to the whole conjunction or in relation to the first conjunct. A CCA pattern emerges only when the conjuncts are reversed, and default agreement is observed:

(84) [*En-a*   *punc-a*   *in*   *pet fant-ov*   ] *je*   *plesal-o.*
     one-F.PL girl(F)-PL and five boy(M)-PL.GEN   AUX.3SG danced-N.SG
     'One girl and five boys danced.'

Finally, if one conjunct is dual and the other a GEN-Q one, the options are: (i) if the dual conjunct is closest, CCA (preferred) or summative number with gender resolution, and (ii) if the GEN-Q conjunct is closest, CCA. However, the preference for CCA if the dual conjunct is closest is available more generally, not just when the other conjunct is a NP with GEN-Q (Šuligoj 2017).

     Although discussing all the intricacies of number agreement with conjoined subjects is impossible here, it is clear that the Slovenian pattern is a highly complex

one. Agreement is sensitive to the number of each conjunct, their case status, and crucially the position of the conjuncts in relation to the agreeing element. Additionally, we saw that interpretative differences might also play an important role in constraining agreement, which is something that is usually not discussed in relation to Slovenian. Thus, although Slovenian has played an important role in the study of CCA, there is plenty more that still needs to be understood about it.

Comitative constructions are another case where, somewhat surprisingly, word order plays a role in number agreement. Slovenian comitative constructions can function as *inclusory conjunctions*: the comitative phrase is interpreted as included in the group denoted by the pronoun it modifies. In other words, 'we with Ana' can mean 'I and Ana' (a cross-linguistically well attested option; see Vassilieva 2005; Vassilieva and Larson 2005; Haspelmath 2007: 33–5; i.a.). In contrast, regular comitative conjuncts, where 'I with Ana' means 'I and Ana', seem to be unavailable (Vassilieva 2005: 72n10). Thus, in (85),[46] unlike the inclusory conjunction with the dual pronoun, the singular pronoun and comitative phrase cannot together bind the reciprocal reflexive, nor can they yield summative dual agreement (dual agreement with the inclusory conjunction is summative inasmuch as the sum of individuals denoted by the conjunction is two).

(85) **Mi-dva**$_{i+k}$ / ***Jaz**$_i$ z   An**-o**$_k$ se$_{i+k}$ *spoštuje-**va**.
    we-DU   / I   with Ana(F)-SG.INS REFL respect-1DU
    'Ana and I respect each other.'

When the pronoun is singular, the comitative phrase is typically an adjunct to the verb phrase, as in (86); note that the comitative phrase is separated from the pronoun, which is impossible with comitative conjuncts (see Vassilieva and Larson 2005: 105).

(86) **Jaz** *pride**-m**   mimo z   An**-o**.
    I   come-1SG by   with Ana(F)-SG.INS
    'I am coming by with Ana.'

What has not been noted before is that summative agreement becomes possible when a singular pronoun and comitative phrase follow the verb, as in (87a), although the comitative phrase must be adjacent to the singular pronoun. Note that these do not behave like true comitative conjunctions: as shown in (87b), they are still incapable of binding reciprocal reflexive when occurring post-verbally.

---

46 Recall that Slovenian is a null-subject language, so in all the examples in (85)–(87), where an overt subject is necessary to distinguish the inclusory and regular conjunctions, the subject is interpreted as discourse-prominent. In comitative constructions with null subjects, it is only the pronoun outside of the comitative phrase that is null.

(87) a. *Pride-va* { ***jaz** } mimo { **jaz** } z An-o.
      come-1DU I by I with Ana(F)-SG.INS
      'I am coming by with Ana.'

   b. *Spoštuje-va* se_{i+k} **mi-dva**$_{i+k}$ / ***jaz**$_i$ z An-o$_k$.
      respect-1DU REFL we-DU / I with Ana(F)-SG.INS
      'Ana and I respect each other.'

These comitative constructions pattern with conjunctions in terms of agreement but not binding. They also show the adjacency requirement unlike verb phrase adjuncts. Lastly, they do not have the relative clause interpretation expected if they were NP-adjuncts (Vassilieva and Larson 2005: 103). It is therefore entirely unclear at this point why they can yield summative agreement.

4 Semantics and Discourse

The interpretation of number in Slovenian has not been discussed thus far, mainly because it is fairly straightforward in the general case. As in other languages with singular, dual, and plural number: singular picks out one, dual two, and plural more than two entities. However, dual is often singled out semantically and pragmatically in Slovenian with respect to singular and plural. For example, unlike singular and plural NPs, bare dual NPs can generally only be used if a co-referring dual NP modified by either 'two' or 'both' has been introduced in the discourse first, as in (88).

(88) *Dv-**a*** *kupc-**a*** *s-ta* *stopil-a* *v* *trgovin-o.*
    two-M.DU buyer(M)-DU AUX-3DU stepped-M.DU in shop(F)-SG.ACC
    'Two shoppers entered the store.

    Ko *s-ta* *si* *kupc-**a*** *ogledal-a blag-o,* *s-**ta***
    when AUX-3DU REFL buyer(M)-DU see-M.DU good(N)-SG.ACC AUX-3DU
    začel-a *naroča-ti.*
    start-M.DU order-INF
    After the shoppers had a look at the goods, they starting ordering.'
    (Jakopin 1966: 99)

Alternatively, the initial dual referent can be introduced through deixis (e.g. with a demonstrative pronoun/phrase or personal pronoun) or by a conjunction with two singular conjuncts. Interestingly, in book/film titles or headlines like (89), bare dual NPs are acceptable without prior introduction of the referent (see Jakopin 1966; Dvořak and Sauerland 2006; Marušič and Žaucer 2021).

(89) *Policist-i   po   Goric-i   lovil-i   tatic-**i**.*
     police.officers(M)-PL around Gorica(F)-LOC chased-PL thief(F)-DU.ACC
     'Police chase (two) thieves around Gorica'
     (Marušič and Žaucer 2021: 440)

These types of differences in the distribution and interpretation of dual as opposed to singular and plural NPs are going to be the main focus of this section. I first discuss some special pragmatic functions of plural and dual (Section 4.1), then I consider the distribution of generic readings and countability (Section 4.2), and finally I consider non-strict readings of plural in Slovenian, including a discussion of the status of pair nouns (Section 4.4).

4.1 Pragmatic functions of number marking

Like many languages, Slovenian uses 2nd person plural for polite/formal addresses. Interestingly, dual is preferred for addressing two individuals for whom a plural form is required individually:

(90) *Kdo ve,   zakaj **vam**   te-ga   ni   povedal-a.*
     Who know.3SG why 2PL.DAT this-SG.ACC not.3SG told-F.SG
     'Who knows why she did not tell you(SG) that.'

     *Saj   s-**ta**   bil-**i**   tako rekoč neločljiv-**i**.*
     PTCL AUX-2DU been.F.DU so   to.say inseparable-F.DU
     'Especially since you two were practically inseparable.'
     (Derganc 2003: 174)

In (90), the first sentence establishes that the addressee must be addressed with a polite 2nd person plural form, however in the second sentence the speaker refers to the group including the addressee by using dual (the subject is omitted, but the agreement is dual) (see also Corbett 2000: 226).

Another relevant phenomenon involving polite forms is mixed number agreement: plural on the verb and singular on the participle, like in (91), used mainly with intermediate levels of formality.

(91) *A bos-**te** še   kaj   pojedl-**a**?*
     Q FUT-2PL still something eat-F.SG
     'Would you still want to eat something?'

Finally, an inclusive 1st person dual may be used in examples like (92), where only the speaker (an acupuncturist) will be performing the action of taking out the needles.

(92) *Gospod Orešnik, zdaj bo-**va** pobral-**a** iglic-e.*
Mr. Orešnik now FUT-1DU take.out-M.DU needle.DIM(F)-PL.ACC
'Mr. Orešnik, we shall take out the needles now.'
(Corbett 2000: 227)

The dual is used here to establish cooperative relationship between the speaker and addressee, but it may also be used to express a patronizing attitude towards the addressee (Derganc 2003: 174–5).

4.2 Genericity

Bare singular and plural NPs can have generic readings in Slovenian, as shown in (93a,b). However, as shown in (93c), dual NPs generally do not allow such readings (Jakopin 1966: 99). Pair nouns are an exception to the lack of generic duals and will be discussed in Section 4.4.

(93) a. *Medved pozimi spi.*
bear(M).SG in.winter sleep-3SG

b. *Medved-**je** pozimi spi-**jo**.*
bear(M)-PL in.winter sleep-3PL
'Bears sleep/hibernate in the winter.'

c. #*Medved-**a** pozimi spi-**ta**.*
bear(M)-DU in.winter sleep-3DU
Only: 'The two bears sleep/hibernate in the winter.'

4.3 Countability

We saw in Section 2.3.3 that Slovenian has several types of number deficient nouns: pluralia tantum nouns, mass nouns, and group/collective nouns – only the latter are truly uncountable.

As noted in the discussion of number deficient nouns above, mass nouns can be counted via measure words or recategorization. Since there are no unambiguously mass determiners/quantifiers in Slovenian (see Section 3.2), one cannot test whether plural nouns can take mass determiners.

Pluralia tantum nouns can also be counted. This can be done via *collective numerals* derived from cardinal ones (Herrity 2000: 137, 2016: 187–90). These numerals behave like true quantifiers (Qs) in that they assign the genitive of quantification (GEN-Q) (see Section 3.2): *dv-**oje** vrat* 'two-DER door(F).PL.GEN' ('two doors'). It

should be noted though that in colloquial speech collective numerals are often substituted with cardinal numerals in this function.

4.4 Non-strict number readings

In Slovenian, both plural and dual can have a weak inclusive interpretation (Marušič and Žaucer 2021), that is: the plural or dual does not strictly have to denote a plurality or duality of entities. This can be observed in questions, under negation, and in if-when contexts (Krifka 1989). For example, in (94) the answers that state that one reviewer was thanked are truthful even though the questions are asking about two (with dual) or more than two reviewers (with plural).

(94) a. *Ali se je avtor zahvalil recenzent-oma? Ja,*
    q REFL AUX.3SG author(M).SG thanked-M.SG reviewer(M)-DU.DAT yes
    en-emu.
    one-M.SG.DAT
    'Did the author thank the two reviewers? Yes, one.'

    b. *Ali se je avtor zahvalil recenzent-om? Ja,*
    q REFL AUX.3SG author(M).SG thanked-M.SG reviewer(M)-PL.DAT yes
    en-emu.
    one-M.SG.DAT
    'Did the author thank the reviewers? Yes, one.'

Additionally, it has been observed that a plural form is sometimes chosen over a dual one even if the referents are clearly two (even outside pair nouns), as in (95).

(95) *T-o je Boeing 737. Motor-je im-a na kril-ih.*
    this-N.SG is Boeing 737 engine(M)-PL.ACC has-3SG on wing(N)-NSG.LOC
    'This is a Boeing 737. Its engines are on the wings.'
    (Dvořak and Sauerland 2006: 105)

To account for this, Dvořak and Sauerland (2006) propose that dual number is purely presuppositional in Slovenian, whereas plural number more generally carries no presuppositions (see also Sauerland 2003; Sauerland, Anderssen, and Yatsushiro 2005). The idea is that the number of the engines is not salient enough in the common ground to satisfy the presupposition of duality, so the plural is used instead; note that (95) concerns the location of the engines on a particular kind of aircraft, not their number (which can only be inferred from real world knowledge about the Boeing 737 or if the engines are visible). Dvořak and Sauerland's analysis also derives the limited distribution of bare dual nouns: since dual number alone cannot be used to assert duality, it merely presupposes it, the use of a bare dual noun is contingent on the prior introduction of a salient dual referent in the discourse. As

we will see below, cases like (95) can help shed some light on the nature of pair nouns.

Pair nouns seem to have a special status when it comes to the interpretation of number, which is relevant if their grammatical status is to be determined. Recall from Section 2.3.3 that pair nouns are a restricted class of nouns in Slovenian that can be plural while referring to a duality – similarly to the plural noun in (95), except the effect is lexically determined. Corbett (2000: 93–4) has argued that pair nouns exemplify the facultative nature of dual in Slovenian, while Derganc (2003: 172–4) has proposed that they are essentially pluralia tantum nouns.

One drawback of Corbett's (2000)'s analysis is that it does not explain why the effect is limited to a small class of nouns. He suggests that the facultative dual is limited by the Animacy Hierarchy: dual is facultative in nouns (low on the hierarchy), but not in pronouns (high on the hierarchy). However, the effect is observed only in a small subset of nouns (mainly paired body parts, items of clothing, and biological roles), and the extent to which specific pair nouns are by default plural or dual varies across dialects. Additionally, animate pair nouns like *starš-i* 'parent(M)-PL' are actually more likely to be plural than inanimate ones like *nog-e* 'leg(F)-PL' (Jakop 2006). Given the Animacy Hierarchy, we would expect more facultative dual in the inanimate nouns.

Since pluralia tantum are lexically constrained, Derganc's (2003) proposal straightforwardly explains at least this aspect of pair nouns. However, recall that all pair nouns can also be dual (e.g. if modified by 'two' or 'both'). Derganc suggests that this is not a regular number alternation: the dual form (e.g. *nog-i* 'leg(F)-DU') is the regular dual of the count noun (*nog-a* 'leg(F)-SG'), while the regular plural form is used to refer to three or more of them (*tri nog-e* 'three.F.PL leg(F)-PL'). The pair noun, as a plurale tantum, is plural in form but not in meaning.

Additional support for Derganc's proposal comes from the fact that pair nouns can be modified by 'two' and remain plural, but only in case 'two' is a collective numeral, as in (96); recall from Section 4.3 that one of the uses of collective numerals is to count pluralia tantum nouns.[47]

(96)   *Dv-oje*   *rok,*   *nog,*   *oč-i,*   *ušes*   *ali*
     two-DER hand(F).PL.GEN leg(F).PL.GEN eye(N).PL.GEN ear(N).PL.GEN or
     pa   *kril.*
     CNTR   wing(F).PL.GEN
     'Two arms, legs, eyes, ears, or wings.'
     (*Gigafida 2.0* corpus: M. Tomšič 2013, *Uroki polne lune III*, novel)

---

47 Related to this, recall that pluralia tantum nouns yield plural agreement on 'one' (Section 3.1; see Corbett 2019: 92 regarding the same fact in Russian). Pair nouns behave the same way: *en-e nog-e* 'one-F.PL leg(F)-PL'. However, it is possible that the plural 'one' is actually related to the set reading: 'one set/pair of legs' (see also Section 3.2). It could be that the two possibilities are connected, but establishing that requires a careful and systematic examination.

Thus, if pair nouns are actually pluralia tantum, they co-exist with a related lexical item which is a regular count noun – something comparable to the count and noun-count uses of 'hair' in English: 'He has a single gray hair' versus 'He has gray hair.' (or the Slovenian *neb-o* 'sky(N)-SG' vs. *nebes-a* 'heaven(N)-PL' from Section 2.3.3). Whether or not such cases are on par with more clear-cut cases of pluralia tantum is not uncontroversial (see Corbett 2019: 63), but regardless of their status, we expect the two variants to have related but nonetheless distinct meanings. In the case of pair nouns and their hypothesized two variants, the relatedness in meaning is uncontroversial. However, the meaning differences are not explored in detail even in Derganc's work.

In order to isolate the meaning difference, we must consider pair nouns not modified by 'two' or 'both'. Crucially, pair nouns can also be dual in such cases, and their meaning is then distinct from the plural form: roughly, the two parts of the pair have a greater degree of independence. A way to illustrate the meaning difference is through the different uses of 'legs' in animals with more than two of them. For example, the passage in (97) describes the way horses walk, and it is crucial how each of the hind legs behaves in relation to the front leg preceding it; although plural is not impossible here, dual conveys the intended meaning more directly.

(97) ... *zato stopat-a zadnj-i nog-i za sledov-i*
      so step-F.DU hind-F.DU leg(F)-DU behind trace(M)-NSG.INS
    prednj-ih nog.
    fore-NSG.GEN leg(F)-NSG.GEN
    '... that is why the hind legs follow in the footprints of the front legs.'
    (*Gigafida 2.0* corpus: 2004, *Revija o konjih*, magazine)

In contrast, the plural is preferred when the behavior or status of each leg is irrelevant, as in (98).

(98) *Vzpenja-jo se na zadnj-e nog-e proti drug-emu*
    rear.up-3PL REFL on hind-F.PL.ACC leg(F)-PL.ACC toward other-N.SG.DAT
    žrebe-tu ...
    foal(N)-SG.DAT
    'They rear up on their hind legs facing the other foal ...'
    (*Gigafida 2.0* corpus: 2005, *Revija o konjih*, magazine)

What makes it difficult to find cases with absolute complementary distribution of the two number variants is that pair nouns as a class encompass exactly those nouns where duality is prototypically part of their lexical semantics (see Jakop 2006, 2008: Ch. 1, 2020: 382 regarding this point, also in relation to the historical development of pair nouns in Slovenian). In the case of 'legs', we normally talk

about them in animals with bilateral symmetry and an even number of them, so the difference between a 'pair of legs' as a unit versus 'legs' as a duality is very rarely informative. However, one can construct contexts where the difference can be controlled for, such as the one in (99).[48]

(99) CONTEXT: *A dead butterfly is examined to assess the effect of insecticides.*
    a. *T-i   nog-i   s-ta   poškropljen-i, ostal-e   pa   ne.* (*dual*)
       this-F.DU leg(F)-F.DU AUX-3DU sprayed-F.DU other-F.PL CNTR not
       'These two legs were sprayed, the others were not.'

    b. *T-e   nog-e   s-o   poškropljen-e, ostal-e pa   ne.* (*plural*)
       this-F.PL leg(F)-F.PL AUX-3PL sprayed-F.PL other-F.PL CNTR not
       'This pair of legs was sprayed, the others were not.'

The two sentences mean different things in (99). When two out of the six legs have insecticide on them, (99a) describes any combination of two legs, but (99b) can only describe the pair of two legs on any of the three segments of the insect. There are therefore scenarios where (99a) is true and (99b) is not. Similarly, plural 'legs' cannot refer to two legs of a radially symmetric seven-legged alien, or two legs of a table with three legs (in case we lead more boring lives).

The presence of a meaning difference fits the pluralia tantum analysis, but the difference is also systematic with pair nouns for each plural-dual pair – it is not specific to 'legs'. This is not necessarily unexpected under the view that pluralia tantum should be subsumed under the broader class of lexical plurals (see Acquaviva 2008: 15–21). In fact, I suggest that this view might even allow us to reconcile the pluralia tantum analysis of pair nouns with the facultative dual one.

Dvořak and Sauerland's (2006) 'Boeing 737 sentence' in (95) shows that plural nouns can refer to dualities outside pair nouns and, consequently, that Corbett's (2000) proposal that dual is facultative in Slovenian is not entirely incorrect, it only defines the contexts where it is facultative too broadly. Dvořak and Sauerland propose that the plural is used in the relevant contexts because the duality presupposition of the dual is not met and a plural meaning is not incompatible with the situation; they even suggest cases like (95) are related to pair nouns, but do not offer an analysis.

Suppose that a lexical item can assert duality by virtue of its lexical semantics. Such a lexical item would be predicted not to require a salient dual referent in the prior discourse. This is exactly what we see with pair nouns, which are natural class comprising of lexemes denoting entities that prototypically come in pairs. Related to this, the meaning of lexical plurals is much more predictable than commonly assumed: their meaning is largely determined based on the denotation of the lex-

---

48 I thank an anonymous reviewer for suggesting this particular scenario.

eme (see Acquaviva 2008: Ch. 4). I tentatively suggest that this is how the meanings of plural and dual pair nouns arise: duality is asserted by the lexeme, but number narrows it down to a specific aspect of plurality/duality. Roughly, plural brings out a collective meaning and dual a distributive one. Generally, the plural form is preferred as it lacks a number presupposition (it is less marked), but in contexts like those described above, each form is picked based on its specialized meaning.

Outside pair nouns, bare dual nouns cannot be used without an already established referent. This explains why bare dual nouns cannot be generic (see Section 4.2), but also predicts that bare pair nouns, which also assert duality just like explicitly quantified dual nouns, should allow a generic reading whether they are plural or dual. As shown in (100), this prediction is borne out.

(100) a. *Pri medved-ih   starš-**i**   skrbij-**jo**   za mladič-e.*
       at bear(M)-NSG.LOC parent(M)-PL take.care-3PL for cub(M)-PL.ACC

   b. *Pri medved-ih   starš-**a**   skrbi-**ta**   za mladič-e.*
      at bear(M)-NSG.LOC parent(M)-DU take.care-3DU for cub(M)-PL.ACC
      'With bears, the parents take care of their cubs.'

The facultative dual and the pluralia tantum analysis of pair nouns are thus in a way both right. Dual is more generally facultative, since plural nouns can refer to dualities outside the constrained lexical class of pair nouns. Nonetheless, the reason why plural number can behave this way is different in the two cases. In the case of pair nouns, it is because the lexeme inherently asserts duality, whereas in all the other cases it is because the dual presupposition is not met.

5 Conclusion

In the domain of grammatical number, Slovenian has thus far mainly received attention due to the rarity of dual number in Indo-European. However, its three-way number system has much more to offer when we closely examine how it interacts with other grammatical categories, for example in the morphological domain in relation to syncretism, suppletion, and derivational morphology.

Even in areas that have already received much attention in Slavic more broadly, such as the syntax of quantified noun phrases and agreement with coordinated phrases, there are many ways in which Slovenian either stands out or offers a new perspective on well studied phenomena.

Finally, the area with perhaps the most potential for future studies is the semantic and pragmatic functions of the Slovenian dual. A closer examination of the semantics of number can help shed light on open puzzles like the status of number deficient nouns, facultative number, and semantic differences between dual and plural beyond the basic 'two' versus 'more than two' opposition.

Acknowledgements

List of Abbreviations

1	1st person
2	2nd person
3	3rd person
SG	singular
PL	plural
DU	dual
NSG	non-singular
M	masculine
F	feminine
N	neuter
NOM	nominative
ACC	accusative
GEN	genitive
LOC	locative
DAT	dative
INS	instrumental
AUG	augment
CNTR	contrastive (topic) particle
DER	derivational morpheme
PTCL	particle
Ø	zero morpheme
LK	linking morpheme
AUX	auxiliary
REFL	reflexive
QMOD	quantity modifiers
NP	Noun Phrase
GEN-Q	genitive of quantification
Q	quantifier
A_Q	quantity adjective
N_Q	quantity noun
Dem	demonstrative

References

Acquaviva, Paolo. 2008. *Lexical Plurals: A Morphosemantic Approach*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. & R. M. W. Dixon. 1998. Dependencies between grammatical systems. *Language* 74. 56–80.

Alexiadou, Artemis, Gianina Iordăchioaia & Elena Soare. 2010a. Number/aspect interactions in the syntax of nominalizations: A Distributed Morphology approach. *Journal of Linguistics* 46. 537–574.

Alexiadou, Artemis, Gianina Iordăchioaia & Elena Soare. 2010b. Plural marking in argument supporting nominalizations. In Patricia Cabredo Hofherr & Brenda Laca (eds.), *Layers of Aspect*, 1–22. Stanford, CA: CSLI Publications.

Ambrazas, Vytautas (ed.). 2006. *Lithuanian Grammar* (2nd edn.). Vilnius: Baltos Lankos.

Anderson, Stephen R. 1969. *West Scandinavian vowel systems and the ordering of phonological rules*. Ph. D. thesis, MIT.

Anderson, Stephen R. 1982. Where's morphology? *Linguistic Inquiry* 13. 571–612.

Anderson, Stephen R. 1986. Disjunctive ordering in inflectional morphology. *Natural Language & Linguistic Theory* 4. 1–32.

Andrews, Avery. 1982. The representation of case in modern Icelandic. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, 427–503. Cambridge, MA: MIT Press.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.

Arsenijević, Boban, Jana Willer-Gold, Nadira Aljović, Nermina Čordalija, Marijana Kresić Vukosav, Nedžad Leko, Frane Malenica, Franc Lanko Marušič, Tanja Milićev, Nataša Milićević, Petra Mišmaš, Ivana Mitić, Anita Peti-Stantić, Branimir Stanković, Jelena Tušek & Andrew Nevins. 2020. Elided clausal conjunction is not the only source of closest-conjunct agreement: A picture-matching study. *Syntax* 23(1). 78–104.

Babby, Leonard H. 1987. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language & Linguistic Theory* 5. 91–138.

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The Syntax-Morphology Interface: A Study of Syncretism*. Cambridge: Cambridge University Press.

Battistella, Edwin L. 1996. *The Logic of Markedness*. Oxford: Oxford University Press.

Bierwisch, Manfred. 1967. Syntactic features in morphology: General problems of so-called pronominal inflection in German. In *To Honor Roman Jakobson: Essays on the Occasion of His 70. Birthday, 11. October 1966*, 239–270. The Hague/Paris: Mouton.

Bittner, Maria & Kenneth Locke Hale. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27(1). 1–68.

Bobaljik, Jonathan David. 2002. Syncretism without paradigms: Remarks on Williams 1981, 1994. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2001*, 53–85. Dordrecht: Kluwer.

Bobaljik, Jonathan David. 2015. Suppletion: Some theoretical implications. *Annual Review of Linguistics* 1. 1–18.

Bonet, Eulalia. 1991. *Morphology after Syntax: Pronominal clitics in Romance languages*. Ph. D. thesis, MIT.

Booij, Geert. 2001. *The morphology of Dutch*. Oxford: Oxford University Press.

Bošković, Željko. 2006. Case and agreement with genitive of quantification in Russian. In Cedric Boeckx (ed.), *Agreement Systems*, 99–120. Amsterdam/Philadelphia: John Benjamins.

Bošković, Željko. 2008. A minimalist account of genitive of quantification. In Luka Szucsich, Gerhild Zybatow, Uwe Junghanns & Roland Meyer (eds.), *Recent issues in formal Slavic*

linguistics. Contributions of the Fifth European Conference on Formal Description of Slavic Languages (FDSL 5), 270–287. Frankfurt am Main: Peter Lang.

Bošković, Željko. 2018. On extraction out of inherently case-marked elements. In Wayles Browne, Miloje Despić, Naomi Enzinna, Simone Harmath de Lemos, Robin Karlin & Draga Zec (eds.), *Formal Approaches to Slavic Linguistics. The Third Cornell Meeting 2016*, 57–78. Ann Arbor, MI: Michigan Slavic Publications.

Caha, Pavel. 2009. *The Nanosyntax of Case*. Ph. D. thesis, University of Tromsø.

Caha, Pavel. 2019. Syncretism in morphology. In Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*, 1–28. Oxford: Oxford University Press. https://oxfordre.com/ linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-552. (accessed 31 January 2021).

Calabrese, Andrea. 1998. Some remarks on the Latin case system and its development in Romance. In José Lema & Esthela Treviño (eds.), *Theoretical Analyses on Romance languages: Selected Papers from the 26th Linguistic Symposium on Romance Languages (LSRL XXVI)*, Mexico City, 28–30 March 1996, 71–126. Amsterdam/Philadelphia: John Benjamins.

Calabrese, Andrea. 2008. On absolute and contextual syncretism: Remarks on the structure of case paradigms and on how to derive them. In Asaf Bacharach & Andrew Nevins (eds.), *Inflectional Identity*, 156–205. Oxford: Oxford University Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, 184–221. Waltham: Ginn.

Chomsky, Noam. 1986. *Knowledge of Language, its Nature, Origin, and Use*. New York: Praeger.

Conklin, Harold C. 1962. Lexicographic treatment of folk taxonomies. In Fred W. Householder & Sol Saporta (eds.), *Problems in lexicography*, 119–141. Bloomington, IN: Indiana University Research Center in Anthropology, Folklore and Linguistics.

Corbett, Greville G. 1979. Adjective movement. *Nottingham Linguistic Circular* 8. 1–10.

Corbett, Greville G. 1983. *Hierachies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croom Helm.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2005. Suppletion in personal pronouns: Theory versus practice, and the place of reproducibility in typology. *Linguistic Typology* 9. 1–23.

Corbett, Greville G. 2019. Pluralia tantum nouns and the theory of features: a typology of nouns with non-canonical number properties. *Morphology* 29. 51–108.

Croft, William. 2003. *Typology and universals* (2nd edn.). Cambridge: Cambridge University Press.

De Belder, Marijke. 2013. Linking phonemes are class markers. Paper presented at TiN-dag 2013, Utrecht, Netherlands, 9th February 2013.

Derganc, Aleksandra. 1994. Some specific features in the development of the dual in Slovene as compared to other Slavic languages. *Linguistica* 34(1). 71–80.

Derganc, Aleksandra. 2003. The dual in Slovenian. In Janez Orešnik & Donald D. Reindl (eds.), *Slovenian from a typological perspective*, 165–181. Berlin: Akademie Verlag.

Derganc, Aleksandra. 2006. Some characteristics of the dual in Slovenian. *Slavistična revija* 54(4). 415–434. Special issue: *Slovensko jezikoslovje danes* [Slovenian linguistics today].

Despić, Miloje. 2017. Investigations on mixed agreement: polite plurals, hybrid nouns and coordinate structures. *Morphology* 27. 253–310.

Dvořak, Boštjan & Uli Sauerland. 2006. The semantics of the Slovenian dual. In James Lavine, Steven Franks, Mila Tasseva-Kurktchieva & Hana Filip (eds.), *Proceedings of FASL 14, The Princeton Meeting*, 98–112. Ann Arbor, MI: Michigan Slavic Publications.

Dziwirek, Katarzyna. 1990. Default agreement in Polish. In Katarzyna Dziwirek, Patrick Farrell & Errapel Mejías Bikandi (eds.), *Grammatical Relations: A Cross-Theoretical Perspective*, 147–161. Stanford, CA: CSLI Publications.

Evans, Nicholas, Dunstan Brown & Greville G. Corbett. 2001. Dalabon pronominal prefixes and the typology of syncretism: A Network Morphology analysis. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2000*, 187–231. Dordrecht: Kluwer.

Falk, Yehuda N. 1991. Case: Abstract and morphological. *Linguistics* 29. 197–230.

Fenger, Paula & Gísli Rúnar Harðarson. 2018. Your *n*s are numbered! On linking morphemes in Dutch. In *Proceedings of the Linguistic Society of America*, Volume 3, paper 26, 1–15. Linguistic Society of America.

Fenger, Paula & Gísli Rúnar Harðarson. 2020. One classy number: Linking morphemes in Dutch and German. In Mariam Asatryan, Yixiao Song & Ayana Whitmal (eds.), *Proceedings of NELS* 50, Volume 2, 1–14. Amherst, MA: GLSA.

Franks, Steven. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language & Linguistic Theory* 12(4). 597–674.

Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. New York: Oxford University Press.

Giannakidou, Anastasia & Melita Stavrou. 1999. Nominalization and ellipsis in the Greek DP. *The Linguistic Review* 16. 295–331.

Giusti, Giuliana & Nedžad Leko. 1995. On the syntax of quantity expressions in Bosnian. *University of Venice Working papers in Linguistics* 5(2). 23–47.

Giusti, Giuliana & Nedžad Leko. 2005. The categorial status of quantity expressions. *Lingvistički Vidici* 34. 121–183.

Golden, Marija & Milena Milojević Sheppard. 2008. The pronominal clitic of quantified noun phrases in Slovenian. Manuscript, University of Ljubljana. http://oddelki.ff.uni-mb.si/filozofija/files/Festschrift/Dunjas_festschrift/goldensheppard.pdf (accessed 31 January 2021).

Golden, Marija & Milena Milojević Sheppard. 2009. The genitive-of-quantification clitic in Slovenian. In Steven Franks, Vrinda Chidambaram & Brian Joseph (eds.), *A Linguist's Linguist: Studies in South Slavic Linguistics in Honor of E. Wayles Brown*, 235–251. Bloomington, IN: Slavica Publishers.

Greenberg, Joseph. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Halle, Morris. 1997. Impoverishment and fission. In Benjamin Bruening (ed.), *PF: Papers at the interface*, Volume 30, 425–450. MIT Working Papers in Linguistics.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser (eds.), *The View from Building* 20, 111–176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2008. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.

Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language Typy and Linguistic Description, Volume II: Complex constructions* (2nd edn.). 1–51. Cambridge: Cambridge University Press.

Herrity, Peter. 2000. *Slovene: A comprehensive grammar*. London/New York: Routledge.

Herrity, Peter. 2016. *Slovene: A comprehensive grammar* (2nd edn.). London/New York: Routledge.

Hopkins, Paul. 2001. *The Phonological Structure of the Kashubian Word*. Ph. D. thesis, State University of New York.

Howe, Stephen. 1996. *The Personal Pronouns in the Germanic Languages*. Berlin/New York: De Gruyter.

Humboldt, Wilhelm von. 1828. *Über den Dualis*. Berlin: Druckerei der Königlichen Akademie der Wissenschaften. [lecture presented at the Berlin Academy of Sciences on 26th April 1827] Reprinted 1907 in: *Wilhelm von Humboldt: Gesammelte Schrifte*, Volume 6: 1827–1835, ed. Albert Leitzmann, 4–30. Berlin: B. Behr's Verlag.

Humboldt, Wilhelm von. 1997. On the dual form. In Theo Harden & Daniel J. Farrelly (eds.), *Essays on language*, 111–136. Frankfurt am Main: Peter Lang.

Jakobson, Roman. 1936. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. *Travaux du Cercle Linguistique de Prague* 6. 240–288. Reprinted in 1971 in: *Selected Writings*, Volume 2, 23–71. The Hague/Paris: Mouton.

Jakobson, Roman. 1958. Morfologičeskije nabljudenija nad slavjanskim skloneniem (sostav russkix padežnyx form). In *American contributions to the IVth International Congress of Slavists*, 127–156. The Hague/Paris: Mouton. Reprinted in 1971 in: *Selected Writings*, Volume 2, 154–181. The Hague/Paris: Mouton.

Jakobson, Roman. 1960. The gender pattern of Russian. *Studii şi Cercetări Lingvistice XI*. 541–543. Reprinted in 1971 in: *Selected Writings*, Volume 2, 184–186. The Hague/Paris: Mouton.

Jakobson, Roman. 1984a. Contribution to the general theory of case: General meanings of the Russian cases. In Linda R. Waugh & Morris Halle (eds.), *Russian and Slavic Grammar*, 59–104. Berlin: Mouton. English translation of Jakobson 1936.

Jakobson, Roman. 1984b. Morphological observations on Slavic declension (the structure of Russian case forms). In Linda R. Waugh & Morris Halle (eds.), *Russian and Slavic Grammar*, 105–134. Berlin: Mouton. English translation of Jakobson 1958.

Jakop, Tjaša. 2006. Nouns in pairs in Slovene dialects. In Svetla Koeva & Mila Dimitrova-Volchanova (eds.), *The Fifth International Conference of Formal Approaches to South Slavic and Balkan Languages* (FASSBL 5): Proceedings, 194–198. Sofia: The Bulgarian Academy of Sciences.

Jakop, Tjaša. 2008. *The dual in Slovene dialects*. Bochum: Universitätsverlag Brockmeyer.

Jakop, Tjaša. 2012. Use of dual in standard Slovene, colloquial Slovene and Slovene dialects. *Linguistica* 52(1). 349–362.

Jakop, Tjaša. 2020. Slovene nouns and adjectives: Non-standard use of dual number from 1925 to 2015. In Stanley D. Brunn & Roland Kehrein (eds.), *Handbook of the Changing World Language Map*, 375–390. Dordrecht: Springer, Cham.

Jakopin, Franc. 1966. Slovenska dvojina in jezikovne plasti [Slovenian dual and language layers]. *Jezik in slovstvo* 11(4). 98–104.

King, Gareth. 2016. *Modern Welsh: A Comprehensive Grammar* (3rd edn.). London/New York: Routledge.

Kiparsky, Paul. 1973. "Elsewhere" in phonology. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 93–106. New York: Holt, Rinehart and Winston.

Klinar, Stanko. 1996. Slovene ali Slovenian – ali kako po Steinbeckovem navdihu ("In dubious battle") bojujemo N/negotovo bitko. *Vestnik DTJK* 30(1/2). 245–253.

Koutsoudas, Andreas. 1968. The A-over-A convention. *Linguistics* 46. 11–20.

Kramer, Ruth. 2016. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry* 47(3). 527–559.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Bentham & Peter van Emde Boas (eds.), *Semantics and contextual expressions*, 75–155. Dordrecht: Foris.

Krott, Andrea, Robert Schreuder & R. Harald Baayen. 2002. Analogical hierarchy: Exemplar-based modeling of linkers in Dutch noun-noun compounds. In Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds.), *Analogical modeling: An exemplar-based based approach to language*, 181–206. Amsterdam/Philadelphia: John Benjamins.

Krott, Andrea, Robert Schreuder, R. Harald Baayen & Wolfgang Uli Dressler. 2007. Analogical effects on linking elements in German compound words. *Language and Cognitive Processes* 22. 25–57.

Lamb, William. 2008. *Scottish Gaelic Speech and Writing: Register Variation in an Endangered Language*. Belfast: Cló Ollscoil na Banriona.

Leko, Nedžad. 1986. *Syntax of noun headed structures in Serbo-Croatian and corresponding phrasal structures in English*. Ph. D. thesis, Indiana University.

Leko, Nedžad. 1989. Case of noun phrases in Serbo-Croatian and Case Theory. *Folia Linguistica* 23(1–2). 27–54.

Lobeck, Anne. 1990. Functional heads as proper governors. In Juli Carter (ed.), *NELS 20: Proceedings of the 20th Annual Meeting of the North East Linguistics Society*, Volume 1, 348–362. Amherst, MA: GLSA.

Lobeck, Anne. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford: Oxford University Press.

Logar, Tine & Jakob Rigler. 1983 [2016]. Karta slovenskih narečij [Linguistic map of Slovenian dialects]. Ljubljana: DDU Univerzum. Updated in 2016 by the Dialectological Section of the Fran Ramovš Institute of the Slovenian Language. https://fran.si/204/sla-slovenski-lingvisticni-atlas/datoteke/SLA_Karta-narecij.pdf (accessed 31 January 2021).

MacAulay, Donald. 1992. *The Celtic Languages*. Cambridge: Cambridge University Press.

Marušič, Franc & Andrew Nevins. 2010. Two types of neuter: Closest-conjunct agreement in the presence of '5&ups'. In Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici, Nikola Predolac & Draga Zec (eds.), *Formal Approaches to Slavic Linguistics. The Second Cornell Meeting 2009*, 301–317. Ann Arbor, MI: Michigan Slavic Publications.

Marušič, Franc, Andrew Nevins & William Badecker. 2015. The grammars of conjunction agreement in Slovenian. *Syntax* 18. 39–77.

Marušič, Franc, & Rok Žaucer. 2021. Dual in Slovenian. In Patricia Cabredo Hofherr & Jenny Doetjes (eds.), The Oxford Handbook of Grammatical Number, 428–444. Oxford: Oxford University Press.

Mel'čuk, Igor. 1994. Suppletion: Toward a logical analysis of the concept. *Studies in Language* 18. 339–410.

Mileusnić, Maja. 2009. *Slovene* ali *Slovenian* [Slovene or Slovenian]. In Jana Zemljaric Miklavčič, Jaka Čibej, Dejan Debelak, Živa Furlan & Manca Hočevar (eds.), *Izzivi jezika: zbornik študentskih strokovnih besedil s področja prevajanja*, 54–57. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta.

Moskal, Beata. 2015a. *Domains at the border: Between morphology and phonology*. Ph. D. thesis, UConn.

Moskal, Beata. 2015b. Limits on allomorphy: A case study in nominal suppletion. *Linguistic Inquiry* 46(2). 363–376.

Müller, Gereon. 2004. On decomposing inflection class features: Syncretism in Russian noun inflection. In Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.), *Explorations in Nominal Inflection*, 189–227. Berlin: Mouton de Gruyter.

Neeleman, Ad & Fred Weerman. 1998. *Flexible Syntax: A Theory of Case and Arguments*. Dordrecht: Kluwer.

Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42(3). 413–444.

Nevins, Andrew & Philipp Weisser. 2019. Closest conjunct agreement. *Annual Review of Linguistics 5*. 219–241.

Noyer, Rolf. 1992. *Features, positions and affixes in autonomous morphological structure*. Ph. D. thesis, MIT.

Noyer, Rolf. 1998. Impoverishment theory and morphosyntactic markedness. In Steven G. Lapointe, Diane K. Brentari & Patrick M. Farrell (eds.), *Morphology and its relation tophonology and syntax*, 264–285. Stanford, CA: CSLI.

Perlmutter, David & Janez Orešnik. 1973. Language particular rules and explanation in syntax. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 419–459. New York: Holt, Rinehart and Winston.

Plank, Frans. 1994. Homonymy vs. suppletion: A riddle (and how it happens to be solved in …). In *Agreement gender number genitive & (EUROTYP Working Papers* VII/23), 81–86. Konstanz: University of Konstanz.

Priestly, Tom. 1993. Slovene. In Greville G. Corbett & Bernard Comrie (eds.), *The Slavonic Languages*, Chapter 8, 388–451. London/New York: Routledge.

Saito, Mamoru & Keiko Murasugi. 1990a. N'-deletion in Japanese. In Javier Ormazabal & Carol Tenny (eds.), *University of Connecticut Working Papers in Linguistics*, Volume 3, 87–107. Storrs, CT: University of Connecticut.

Saito, Mamoru & Keiko Murasugi. 1990b. N'-deletion in Japanese: A preliminary study. In Hajime Hoji (ed.), *Japanese/Korean Linguistics*, Volume 1, 285–302. Stanford, CA: CSLI.

Sauerland, Uli. 2003. A new semantics for number. In Robert Young & Yuping Zhou (eds.), *Proceedings of SALT* 13, 258–275. Ithaca, NY: CLC Publications.

Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stefan Kepser & Marga Reis (eds.), *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*, 409–430. Berlin: De Gruyter.

Schenker, Alexander M. 1993. Proto-Slavonic. In Greville G. Corbett & Bernard Comrie (eds.), *The Slavonic Languages*, Chapter 3, 60–124. London/New York: Routledge.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer.

Shen, Zheng. 2018. *Feature Arithmetic in the Nominal Domain*. Ph. D. thesis, UConn.

Slobodchikoff, Tatyana G. 2013. *The Slavic Dual: Morphosyntactic Feature Economy as a Factor in Language Change*. Ph. D. thesis, University of Arizona.

Slobodchikoff, Tatyana G. 2019. *The Evolution of the Slavic Dual: A Biolinguistic Perspective*. London: Lexington Books.

Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2019. Case and number suppletion in pronouns. *Natural Language & Linguistic Theory* 37. 1029–1101.

Starke, Michal. 2001. *Move reduces to Merge: A theory of locality*. Ph. D. thesis, University of Geneva.

Stegovec, Adrian. 2016. Not two sides of one coin: Clitic person restrictions and Icelandic quirky agreement. In Lanko Marušič & Rok Žaucer (eds.), *Formal studies in Slovenian Syntax. In Honor of Janez Orešnik*, 283–312. Amsterdam: John Benjamins.

Stegovec, Adrian. 2020. Hidden deficiency: On the structure of slovenian clitic, strong, and prepositional pronouns. Paper presented at the 15th Annual Meeting of the Slavic Linguistics Sociaty (SLS), Indiana University, September 3rd–6th. https://slavicgf.sitehost.iu.edu/SLS/slides/Stegovec_SLS_pronouns_ho.pdf (accessed 31 January 2021).

Stone, Gerald. 1993. Cassubian. In Greville G. Corbett & Bernard Comrie (eds.), *The Slavonic Languages*, Chapter 13, 759–794. London/New York: Routledge.

Stump, Gregory T. 1993. On rules of referral. *Language* 69(3). 449–479.

Stump, Gregory T. 2002. Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2001*, 147–180. Dordrecht: Kluwer.

Stump, Gregory T. 2006. Heteroclisis and paradigm linkage. *Language* 82(2). 279–322.

Šuligoj, Tina. 2017. Slovnična kategorija števila pri stavčnem ujemanju: Dvojina v koordinaciji [The grammatical category of number in clausal agreement: Dual in coordination]. Master's thesis, Univerza v Novi Gorici.

Tesnière, Lucien. 1925a. *Atlas linguistique pour servir à l'étude du duel en slovéne*. Paris: Librarie ancienne Honoré Champion.

Tesnière, Lucien. 1925b. *Les formes du duel en slovène*. Paris: Librarie ancienne Honoré Champion.

Tiersma, Peter. 1982. Local and general markedness. *Language* 58(4). 832–849.

Tollefson, James William. 1978. *Diglossia and language policy, with special reference to Slovenia*. Ph. D. thesis, Stanford University.

Toporišič, Jože et al. (eds.). 2001. *Slovenski Pravopis* [Slovenian Orthography]. Ljubljana: Založba ZRC. Available at: http://www.fran.si/.

Toporišič, Jože. 2000. *Slovenska slovnica* [Slovenian Grammar] (4th edn.). Maribor: Obzorja.

Vassilieva, Masha & Rickard K. Larson. 2005. The semantics of the plural pronoun construction. *Natural Language Semantics* 13. 101–124.

Vassilieva, Maria Borisovna. 2005. *Associative and Pronominal Plurality*. Ph. D. thesis, Stony Brook University.

Veselovská, Lida. 1995. *Phrasal Movement and X⁰ morphology. Word order parallels in Czech and English nominal and verbal projections*. Ph. D. thesis, Palacky University, Olomouc.

Wiltschko, Martina. 2014. *The Universal Structure of categories: Towards a Formal Typology*. Cambridge: Cambridge University Press.

Zwicky, Arnold M. 1985. How to describe inflection. In Mary Niepokuj et al. (ed.), *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, 372–386. Berkeley, CA, Berkeley Linguistics Society.

Irina Bagirokova, Yury Lander, and Paul Phelan

8 Number in West Circassian

Abstract: This chapter describes the expression of number in West Circassian, a polysynthetic language of the Northwest Caucasian family. In West Circassian, nouns contrast between non-specific deficient forms undefined for number, unmarked forms which usually – but not always – express the singular meaning, and forms which explicitly mark plurality (either by a simple plural affix or by a cumulative case-number suffix). Notably, the plural morphology attaches both to count nouns and to mass nouns, sometimes triggering various semantic shifts. The contexts where a form without a plural marker may have plural denotation include not only nominals with non-specific reference but also possessives and non-locutor definite pronominal phrases. Number can also be expressed by a dedicated associative plural suffix (whose use is however severely restricted) and by indexing outside the noun whose use has a predominantly semantic basis. The language further displays some minor constructions, such as the honorific use of indexing, a dyadic construction based on reciprocal morphology, and a pattern close to inclusory constructions.

1 Overview

West Circassian, also known as Adyghe, belongs to the West Caucasian (or Northwest Caucasian / Abkhaz-Adyghe) family – together with Kabardian, its closest relative with whom it constitutes the Circassian branch, and more distantly related Abkhaz, Abaza, and the now extinct Ubykh. The language was originally spoken in the West Caucasus (now part of Russia) but is also used in the Circassian diaspora primarily in Turkey, Syria, Jordan and Israel. The main typologically relevant grammatical descriptions of West Circassian include Rogava & Kerasheva 1966 and Testelets (ed.) 2009 (in Russian), Smeets 1984 (in English) and Paris 1989 (in French), but they are based on different idioms. This chapter describes the expression of number in idioms closest to the standard variant of the language as presented in Russia and used in media and fiction. Our data come mainly from the West Circassian electronic corpus (Arkhangelskiy et al. 2018), which covers various texts in the written standard language, but we also use examples obtained through elicitation (the latter are marked by [E]; we also use square brackets inside the examples to indicate elicitation of examples resulting from modification of corpus examples by adding material).

West Circassian is left-branching (it has a predominantly predicate-final word order, postpositions and possessors preceding their possessa) and ergative both in head- and dependent-marking. Like other members of the family, West Circassian

Tab. 1: Cross-reference markers (basic allomorphs).

	ABS	IO (including ALIENABLE POSSESSORS)	ERG, INALIENABLE POSSESSORS
1SG	*sə-*	*s-*	
1PL	*tə-*	*t-*	
2SG	*wə-*	*p-/w-*	
2PL	*ŝʷə-*	*ŝʷ-*	
3SG	–	*Ø-*	*ə-/ jə-*
3PL	–	*a-*	
REC	–	*ze-*	*ze.re-*
RFL	*zə-*	*zə-*	–
REL	–	*zə-*	

shows many traits which are typically associated with polysynthesis (Lander & Tes-
telets 2017). The predicate may display complex morphology including both prefixa-
tion and suffixation (cf. Smeets 1984). The participants of the event described in a
clause are normally indexed (cross-referenced) by personal prefixes:

– the arguments of the predicate are indexed on the predicate;
– the referential possessor is indexed on the possessum; alienable and inalienable
  possessive constructions are contrasted by the absence/presence of the posses-
  sive prefix *jə-* as well as by the choice of the prefixal series (the indirect object
  series used for arguments introduced by specific morphology vs. a series identi-
  cal to ergative indexing except for the reciprocal marking);
– the object of a postposition is normally (but not always) indexed on the postpo-
  sition.

The inventory of indexing prefixes (except for the indexing of postpositional ob-
jects) is presented in Table 1.[1] The series include reflexive and relative prefixes,
which cannot be linked to any overt external phrases.[2]

---

1 The table includes both zero indexing and the absence of indexing. Some arguments for distin-
guishing between these two treatments for different arguments are presented in Arkadiev et al.
2009 and Lander et al. 2021.
2 The reciprocal prefixes are clearly related to and occasionally are replaced with reflexive ones
(cf. Letuchiy 2007; Ershova 2019), and it is possible that the reflexive and the relative may be two
sides of the same exponent (cf. Daniel & Lander 2019). For the sake of clarity, however, we will
continue to distinguish between these three functions in this chapter.

An example containing the predicate indexing the absolute argument, the ergative argument and an indirect object as well as the possessum indexing the possessor is given below:[3]

(1) *ne-ŝʷə-č̣'e*   *w-a-r-j-e-ʁe-pλ̣ə*
eye-good-INS   2SG.ABS-3PL.IO-DAT-3SG.ERG-DYN-CAUS-look
with.good.eyes it.makes.you.look.at.them
t-jə-nəbž'əč̣'e-xe-r   *dax-ew*   *qə-ze.re-ŝʷe-xe-re-m*
1PL.IO-POSS-youth-PL-ABS beautiful-ADV DIR-REL.FACT-dance-PL-DYN-OBL
our.youth   beautifully   the.fact.that.they.dance
'The fact that our young people dance beautifully makes you admire them
(lit., look at them with good eyes).'

Besides indexing, West Circassian also employs dependent-marking means including postpositions and cases. Traditionally, four cases are distinguished (see, e.g., Rogava & Kerasheva 1966; Smeets 1984; Arkadiev et al. 2009 *inter alia*), here called absolute (used mainly for 3rd person intransitive subjects and transitive undergoers), oblique (used for various cross-referenced roles including the transitive agent), instrumental, and adverbial, but the case status of the latter two is debatable. We will refer to absolute and oblique as core cases. In (1), the oblique case marks a subordinate clause ('that our young people dance beautifully'), the absolute case marks its subject ('our young people'), and the instrumental appears on the adjunct of the matrix predicate ('with good eyes'). Core case markers are usually absent on some pronouns, proper names, possessive phrases and non-specific phrases. Arkadiev & Testelets (2019) demonstrated that non-specific phrases unmarked for case also display deficiency in some other grammatically relevant respects, hence we will call such non-specific NPs *deficient nominals* and contrast them with *unmarked NPs*, by which we understand NPs formally unmarked for number but not for case.

West Circassian demonstrates a very flexible parts-of-speech system. All content words may function as predicates by taking the relevant predicate morphology (e.g., tense and the corresponding indexing prefixes), and as heads of NPs by taking the relevant argument morphology. Nouns can be distinguished as a specific class which combines with relative clause attributes. Further, the combination of a noun (or an adjective) with some of its attributes constitutes a peculiar formation called the *nominal complex*, which has morphonological and grammatical properties of a single word (Lander 2017). Prefixes (e.g., possessive morphology) appear at its very beginning and inflectional suffixes (including the plural and case markers) are found at its very end:

---

3 The simplified word-by-word translation provided in the examples so as to make the structure more obvious is not intended to convey literal semantics or reflect all grammatical features.

(2) **t-jə-čəg-λepq-deʁʷə-xe-m**   *ja-qebar*   *ze.re-dwənaj-ew*
1PL.IO-POSS-tree-kind-good-PL-OBL 3PL.IO+POSS-news ALL-world-ADV
our.good.sorts.of.trees   their.news   throughout.the.world
š'ə-ze-λ-a-ṣe
LOC-REC.IO-LOC-3PL.ERG-know
they.know.it
'Our good sorts of trees are known throughout the world (lit., they know the news of our good sorts of trees throughout the world).'

A noun or a nominal complex may be modified by demonstratives, relative clauses, possessor NPs, and some other attributes.

The information on number may be conveyed by various means in nominals and in non-nominal predicates: by number marking proper (sometimes combined with case marking) and by indexes, by prefixes and by suffixes.

2 Pronominal, nominal, and verbal number

2.1 Generalities

While considering the expression of number in West Circassian nominals, one should distinguish between at least three different systems, namely:
(i)   the number contrasts in personal pronouns,
(i)   the contrasts related to the expression of the additive plural,
(i)   the associative plural.

2.2 Pronominal number

Personal pronouns in West Circassian distinguish two persons (1st vs 2nd) and two numbers (singular vs plural). These distinctions are expressed as different stems (Table 2).

Tab. 2: Personal pronouns.

	Singular	Plural
1st person	*se*	*te*
2nd person	*we*	*ŝʷe*

Evidence that the number contrast in personal pronouns reflects a category which is different from the number category in other parts of grammar comes from the fact

that plural personal pronouns combine with numerals, while the combinability of the (additive) nominal number with numerals is much more restricted (Section 2.3.2):

(3) *səd-a* ***t-jə-ṭʷ-jə*** *zə-čʼa-p-ʔʷe-re-r,* ***t-jə-šʼ***
what-Q we-LNK-two-ADD REL.IO-RSN-2SG.ERG-say-DYN-ABS we-LNK-three
what we.both the.reason.why.you.say we.three
ʔʷe-ba!
say-EMP
say
'Why do you say, "we both"? Say "we three"!'

Personal pronouns almost never receive overt core case marking (except when they appear as objects of some postpositions requiring the oblique case suffix *-šʼ/-j*).

There are no 3rd person pronouns proper in West Circassian. The anaphoric function is fulfilled by the pronoun *a-*, which we gloss below as 'that' as it formally belongs to demonstratives.

Pronouns other than those of the 1st and 2nd persons share the number system with other nominals and hence are described in Section 2.3.

2.3 Nominal number

When speaking of nominal number, we cover all kinds of NPs with the exception of those expressed by personal pronouns. Note that in multiword NPs, the number is usually only marked on the head word, which can, however, be represented by a nominal complex. In (4) we find a combination of a demonstrative, a relative clause, and a noun, but only the noun takes the plural suffix.

(4) [*mə* *sə-zə-ʁe-gʷəmeč̣ʼə-re* ***ʔʷefəʁʷe-xe-m***]
this 1SG.ABS-REL.ERG-CAUS-worry-DYN issue-PL-OBL
these bothering.me issues
ja-ze.ŝʷe.xə-n *ʔʷef-qə.ze.rə.kʷ-ew* *šʼə-t-ep*
3PL.IO+POSS-manage-MOD work-simple-ADV LOC-stand-NEG
solving.them as.an.easy.task it.does.not.stand
'Solving these problems which are bothering me is not an easy task.'

In Sections 2.3.1–2.3.4, we discuss the contrast between deficient forms, unmarked forms and additive plural forms, while in Section 2.3.5 we focus on the associative plural.

2.3.1 Exponence of number

In nominal NPs considered in this section, the basic contrast is that between deficient nominals (5), unmarked nominals (6), and the additive plural forms marked either by the suffix *-xe* or by the oblique case plural suffix *-me* or by both (7).[4] As is seen from the translation and discussed in details below, these three classes differ in number interpretations they can acquire.

(5) *gʷəš'əʔe.wəxəʁ je-pλ!*
sentence   DAT-look
'Look at any sentence(s)!' [E]

(6) *gʷəš'əʔe.wəxəʁe-**m** je-pλ!*
sentence-OBL   DAT-look
'Look at a/the sentence!' [E]

(7) *gʷəš'əʔe.wəxəʁe-**xe-m*** / *gʷəš'əʔe.wəxəʁe-**me*** / *gʷəš'əʔe.wəxəʁe-**xe-me***
sentence-PL-OBL   sentence-OBL.PL   sentence-PL-OBL.PL
ja-pλ!
3PL.IO+DAT-look
'Look at (the/some) sentences!' [E]

The paradigm of the noun *təʁʷəẑ* 'wolf' including case/number forms and the form used in the predicate position (which may take additional morphology) is provided in Table 3. As the table shows, for adverbial and predicative forms, the contrast between deficient and unmarked nominals is irrelevant. The instrumental suffix, probably a former postposition, can be added either to bare stems (for non-specific expressions) or to the stems with an oblique marker *-m* (but not to the oblique plural marker *-me*), so the distinction between the deficient and the unmarked form is retained for instrumental forms.

    The most intriguing part of the paradigm is the oblique plural forms (cf. a detailed discussion in Arkadiev 2014: 555–559). In most contexts we do not find any differences between the three ways of marking the oblique plural, the clearest exception being contexts with numerals (see (96) below). Arkadiev (2014: 556) hypothesizes that the use of the dedicated oblique plural suffix may be more typical of animate NPs and Smeets (1992: 304) reports the suggestions of some speakers that *-xe* (unlike *-me*) implies the reference to the whole set, yet we are not aware of any investigation of these issues. Note also that the oblique plural suffix does not appear in instrumental forms. Given that, in oblique forms, the distribution of the two suffixes is not

---

4 Kerasheva (1982) proposed that the plural value of the number category can also be expressed by *be* 'many'. We do not find evidence for its grammaticalization, even though in some contexts the original evaluation of the size of the set expressed by *be* may indeed be weakened.

Tab. 3: Basic noun paradigm.

	Deficient	Unmarked	Plural
Absolutive	təʁʷəẑ	təʁʷəẑə-r	təʁʷəẑ-xe-r
Oblique		təʁʷəẑə-m	təʁʷəẑ-xe-m təʁʷəẑ-me təʁʷəẑ-xe-me
Instrumental	təʁʷəẑə-ç'e	təʁʷəẑə-m-ç'e	təʁʷəẑ-xe-(m)-ç'e
Adverbial	təʁʷəẑ-ew		təʁʷəẑ-x-ew
Predicative	-təʁʷəẑ-		-təʁʷəẑə-x(e)-

straightforward and their contribution to the interpretation of the number value is potentially different, we often provide illustrations for both constructions.

As mentioned in Section 1, the plural morphology appears in the end of the nominal complex as a whole and hence need not attach to the noun (cf. 'our sorts of trees' in (2) above). Further, we find it with some dvandva compounds optionally marked with a plural suffix such as *jane-jate-me* [POSS-mother-POSS-father-OBL.PL] 'his/her parents' or *šə-šəpχʷə-xe-r* [brother-sister-PL-ABS] 'brother(s) and sister(s)'.

In the standard register discussed here (but not in non-standard varieties; cf. Arkadiev & Testelets 2019: 732), the plural suffix should be accompanied by a case marker except for two contexts: vocative (as in (8); cf. Kumakhov 1971: 5) and an appositive construction containing a proper name following a plural nominal, as in (9), where the plural expression 'the chiefs of the Bahri and the Burji' is followed by proper names 'Salar and Jangashir'.

(8) **pŝaŝe-xe**, ŝʷə-q-jə-ha-ž'!
girl-PL 2PL.ABS-DIR-LOC-go.in-RE
girls go.in.back
'Girls, go home!' [E]

(9) *baχrjət-xe-m-re* *bwərž'jət-xe-m-re* **ja-paš'e-xe** *salarə-re*
Bahri-PL-OBL-COORD Burji-PL-OBL-COORD 3PL.IO+POSS-chief-PL Salar-COORD
Bahri Burji their.chiefs Salar
ž'anšagjərə-re *xebza-ŝhe-r* *ja-ze.re-mə-ʁe-gʷeš'ə*
Jangashir-COORD order-head-ABS 3PL.IO+DAT-REC-ERG-NEG-CAUS-divide
Jangashir leadership they.do.not.share
z-e-χʷə-m …
REL.TMP-DYN-happen-OBL
when.it.happened
'When it so happened that the heads of the Bahri (Mamluks) and the Burji (Mamluks), Salar and Jangashir, did not share power …'

Tab. 4: Basic paradigm of the proximate demonstrative pronoun.

	Unmarked	**Plural**
Absolutive	*mə-r*	*mə-xe-r*
Oblique	*mə-š' / mə-j*	*mə-xe-m*
		mə-xe-me
Instrumental	*mə(-š')-ç'e/*	*mə-xe-m-ç'e*
	mə(-j)-ç'e	
Adverbial	*mə-r-ew*	*mə(-rə)-x-ew*
Predicative	*mə-rə-*	*mə-rə-x(e)-*

A paradigm similar to that presented in Table 3 is found for most non-personal pronouns with the exception of demonstratives. When demonstratives constitute an NP themselves (i.e. do not appear as modifiers), they appear to be similar to personal pronouns in taking a distinct suffix for oblique (but, unlike personal pronouns, not only with postpositions) and in having an additional "predicative" form marked with the suffix *-r(ə)*, which also appears in non-core cases. Cf. the paradigm of the proximate demonstrative *mə* in Table 4.

2.3.2 Core number contrasts

The core number contrast is that between singular and plural. As explained below, the mapping of these two values on the three forms – deficient nominals, unmarked nominals and the plural form – is not trivial, primarily because of the complex status of the unmarked forms.

Deficient nominals are undefined for the number value (see also Arkadiev & Testelets 2019: 731–732). In (10) the use of the bare form of the nominal may be understood to convey any number, and the answer to the question does not provide any more clarity: the speaker merely affirms that there is at least one plum present, but there may be more than just that one. Similarly, a clause such as (11) may describe an event containing a single book or more: according to informal comments made by speakers, 'book' is used here primarily to describe the type of event rather than to make a statement concerning any specific book(s).

(10)   ***qəpçe*** *wəne-m*   *jə-λ-a?*   – *a-rə*
     plum   house-OBL LOC-lie-Q   that-PRED
     plum   house   they.lie.in.it   yes
     'Is there a plum / Are there plums at home? – Yes.' [E]

(11) *s-šə*　　　　　　**txəλ-pəwət**　*qə-s-jə-tə-ʁ*
　　　1SG.PR-brother　book-cheap　DIR-1SG.IO-3SG.ERG-give-PST
　　　my.brother　　　cheap.book　s/he.gave.it.to.me
　　　'My brother gave me a cheap book / cheap books as a gift.' [E]

Plural forms almost always imply that the number of referents is more than one. In particular, there are no plural nouns that can refer to single entities like *scissors* in English. Still, there are certain non-standard uses of plural forms with mass nouns which do not allow a description in terms of simple plurality; see Section 2.3.3.

　　Unmarked forms have been described in literature as if they can refer to single individuals, have plural denotation, or remain undefined for number (cf. Kumakhov 1969; 1971: 6–14; see also Tutarisheva & Tutarisheva 2015); they thus supposedly instantiate "general number" forms, i.e. forms not belonging to the number system (Corbett 2000: 10). Yet, the probability of the use of an unmarked form with plural denotation depends on the register and the context. Apart from contexts where number is undefined (as in generic sentences), at least for the varieties spoken in Russia, in written texts the unmarked form is normally interpreted as singular. In general, the number interpretation in West Circassian seems to depend partially on the referential properties of the nominal; cf. Section 4.1.

2.3.3 Non-standard use of plural form

In this section, we describe some semantic shifts observed in forms containing plural morphology (see also Ilyina 2011). Most of them involve mass and abstract nouns, but at the end of the section we touch upon some other uses.

　　Mass nouns, upon receiving plural marking, can express a range of concepts with the general semantics of multiple units:

(12) *twəč'anə-m fəgʷə-r*　　**ŝʷewš'əʁʷə-me** *a-xe-λə-ʁ*
　　　store-OBL　　millet-ABS　sugar-OBL.PL　3PL.IO-LOC-lie-PST
　　　store　　　　millet　　　sugar(pl)　　　it.lies.among.them
　　　'In the store millet lay among the packs of sugar.' [E]

(13) *depq-xe-m*　　*a-tje-λə-ʁe*　　　　**jaţe-xe-r**　*č'ərbəš'ə-m fe-d-ew*
　　　wall-PL-OBL　3PL.IO-LOC-lie-PST　clay-PL-ABS　brick-OBL　BEN-correspond-ADV
　　　walls　　　　which.lied.on.them　clay(pl)　　brick　　　like
　　　a-tje-ẑəhe-ʁa-ʁe-x …
　　　3PL.IO-LOC-harden-PST-PST-PL
　　　they.hardened.on.them
　　　'The pieces of clay which were left on the walls hardened on them like bricks …'

(14) **dəŝe-xe-r**   zə-š'-a-š'e-re   twəč'an-ew t-jə-qele-ŝhaʔe
      gold-PL-ABS REL.IO-LOC-3PL.ERG-sell-DYN store-ADV   1PL.IO-POSS-city-main
      gold(pl)   where.they.sell.(them)   store   our.capital
      de-t   gʷere-m mə mafe-xe-m je-təʁʷa-ʁe-x
      LOC-stand   some-OBL this day-PL-OBL DAT-rob-PST-PL
      which.stands.there some   this day   was.robbed
      'A certain store which sells gold jewelry and is located in the capital was
      recently robbed.'

Another possible meaning derived from a pluralized mass noun is that of a large
amount of the substance.

(15) *swəlţan psənč'-ew ə-negʷ*   **λə-xe-r**   *š'-jə-f-jə*
      sultan   quick-ADV 3SG.PR-face blood-PL-ABS LOC-3SG.ERG-spread-ADD
      sultan   quickly   his.face   blood(pl)   he.spread.(them).there
      z-jə-ʁe-λa-ʁ
      RFL.ABS-3SG.ERG-CAUS-die-PST
      he.made.himself.die
      'The sultan quickly rubbed a lot of blood all over his face and played dead.'

(16) **ŝʷ-jə-wes-me**   *w-a-fe-zeš'ə-ʁ-a?*
      2PL.IO-POSS-snow-OBL.PL 2SG.ABS-3PL.IO-BEN-miss-PST-Q
      your.snow(pl)   you.missed.them
      'Do you miss your heavy snows?'

At other times, context can demonstrate that different sorts of a material or sub-
stance are being discussed (17). The "multiple sorts" reading is akin to the taxonom-
ic reading observed with count nouns (see example (72) below). With mass nouns,
however, the boundary between different, individual objects and different sorts of
a substance can sometimes be difficult to tell apart, as is the case in (18).[5]

(17) *vəstavke-m   jə-dəŝe-mjedal'-xe-r*   **qʷaje-x-ew**   "adəg'ejsk'ij",
      exhibition-OBL POSS-gold-medal-PL-ABS cheese-PL-ADV Adyghean
      exhibition   its.gold.medals   cheese(pl)   Adyghean
      "adəg'ejsk'ij kopč'onəj" əč'jə "g'iag'insk'ij" "spag'ett'i"
      Adyghean   smoked   and Giaginsky   Spaghetti
      Adyghean   smoked   and Giaginsky   Spaghetti

---

5 The use of numerals with mass nouns results in a similar semantic shift (cf. Nikolaeva 2012).

zə-f-jə-ʔʷe-xe-re-m   *a-paje*
REL.IO-BEN-3SG.ERG-say-PL-DYN-OBL   3PL.IO-for
on.which.he.says   for.them
a-r-a-tə-ž'ə-ʁe-x
3PL.IO-DAT-3PL.ERG-give-RE-PST-PL
they.gave.them.to.them
'At the exhibition they were given gold medals for cheeses with the names
"Adyghean," "Smoked Adyghean," and "Giaginsky Spaghetti".'

(18)   ... *kʷəpə-r*   *zə-š'ə-ze.re-waʁʷejə-re-m*   **lə-xe-r**
      group-ABS   REL.TMP-LOC-REC.ERG-gather-DYN-OBL   meat-PL-ABS
      group   when.(they).gather   meat(pl)
š'-a-wəpš'erəhe-x ...
LOC-3PL.ERG-cook-PL
they.cook.them.there
'... when they gathered as a group, they cook various meat dishes ...'

Abstract parameter nouns may also be pluralized. In these contexts, the abstract
noun is considered to have numerous possessors and/or imply multiple manifesta-
tions of the relevant properties, emotions etc. Cf. examples (19)–(20) which involve
multiple possessors and (21)–(22) without any possessive morphology. The last of
these examples demonstrates that the plural suffix may appear on action nominals
(derived by the so-called "modal" suffix).

(19)   **ja-ç'ʔəhaʁe-xe-m-ç'e**   *mjetre-pŝə~rəpŝ*   *χʷə-š'tə-ʁe-x* ...
     3PL.IO+POSS-length-PL-OBL-INS   meter-ten~DISTR   happen-AUX-PST-PL
     by.their.length(pl)   ten.meters   they.happened
     'Their lengths were (about) ten meters ...'

(20)   *jəljase-re*   *dexezač'e-re*   **ja-ŝʷəλeʁʷənəʁe-xe-r**
     Ilyas-COORD   Dakhazache-COORD   3PL.IO+POSS-love-PL-ABS
     Ilyas   Dakhazache   their.love(pl)
     zə-ze-pe-p-š'eč'ə-ç'e ...
     REL.TMP-REC.IO-LOC-2SG.ERG-weigh-INS
     if.you.weigh.(them)
     'If you are to assess Ilyas's and Dakhazache's love ...'

(21)   *jəλes-ew*   *jə-ç'ə-re-m*   *nah*   *qə-xe-b-ʁe.š'ə-n-ew*
     year-ADV   LOC-go.out-DYN-OBL   more   DIR-LOC-2SG.ERG-reveal-MOD-ADV
     year   last   more   you.reveal.(him).there
     gʷəŝʷaʁʷe-xe-r *qə-xe-fa-ʁe-x*
     joy-PL-ABS   DIR-LOC-fall-PST-PL
     joy(pl)   they.fell.there
     'Last year there were joys which are worth revealing.'

(22) ... ***gʷəbžə-n-xe-r***,   ***gʷəmeč'ə-n-xe-r***   *ṭekʷ~ṭekʷ-ew*   *nah*   *mač'e*
     be.angry-MOD-PL-ABS   worry-MOD-PL-ABS   gradually-ADV   more   little
     anger   worry   gradually   more   little
me-χʷə ...
DYN-happen
(they).become
'Anger and worry are gradually decreasing ...'

Toponyms, if modified with the plural, indicate a number of denizens of that place (23), although this is more typical for well-established place names and is likely to be related to the occasional appearance of toponyms referring to individual denizens (not illustrated here). Less typically, group nouns or collective nouns are pluralized to express plurality of the group's participants (24)–(25).

(23) ***wəlape-me***   *ap-ew*   *qelapče-m*   *ʔegʷawe-r*   *d-a-ʒa-ʁ*
     Ulyap-OBL.PL   first-ADV   goal-OBL   ball-ABS   LOC-3PL.ERG-kick-PST
     Ulyaps   first   goal   ball   they.kicked.it.into.it
     'The Ulyaps first kicked the ball into the goal.'

(24) *sə-z-e-səmaž'e-m*   ***s-jə-klass-xe-r***   *mafe-qes*
     1SG.ABS-REL.TMP-DYN-sick-OBL   1SG.IO-POSS-class-PL-ABS   day-every
     when.I.became.sick   my.class(pl)   every.day
qə-s-fa-ḳʷe-š'tə-ʁe-x
DIR-1SG.IO-BEN-go-AUX-PST-PL
they.visited.me
'When I was sick my classmates visited me every day.' [E]

(25) ***m'il'icije-xe-m*** *q-a-ʁe-wəcʷə-ʁe*   *mašəne-m*   *sp'irt*
     militia-PL-OBL   DIR-3PL.ERG-CAUS-stand.up-PST   car-OBL   spirit
     militia(pl)   which.they.put   car   spirit
ze-rə-t   *bešereb-ljətre-nəqʷe*   *2300-re*   *qə-r-a-ʁʷeta-ʁ*
REL.IO-LOC-stand   bottle-liter-half   2300-COORD   DIR-LOC-3PL.ERG-find-PST
where.it.stands   half-liter.bottle   2300   they.found.there
'In the car left by the militia officers they found 2300 half-liter bottles of spirit.'

While the plural suffix *-xe* can create family names in the closely related Kabardian (Kumakhov (ed.) 2006: 89), in West Circassian it usually conveys nothing more than plurality. Thus, in (26) we find a family name *asλan(e)* which coincides with a widespread first name: still, it is used without any additional suffix in the first of these examples and only takes the plural marker when the NP has plural denotation, as in the second example.

(26) a. **asɬane** *tjəmwər t-jə-neʔʷas*
Aslan   Timur   1PL.IO-POSS-acquaintance
Aslan   Timur   he.is.our.acquaintance
'Aslan [family name] Timur is our acquaintance.' [E]

    b. **asɬan-xe-r** *nahə-b-ew   grjekə-x*
Aslan-PL-ABS more-many-ADV Greek-PL
Aslan   more   they.are.Greeks
'The Aslans [family name] are mostly Greeks.' [E]

2.3.4 Reciprocal dyadic plural

An interesting pattern which represents a kind of dyadic construction (Evans 2006) is found with symmetric relational nouns. If their plural forms refer to a closed set whose members stand in a relation denoted by the root, these forms take a reciprocal prefix *ze-* (cf. Letuchiy 2007: 800):

(27) *qare* **ze-šə-me**   *a-nahə-ẑə-r*   *č'ela.č'-ew λa-ʁe*
Kare REC.PR-brother-OBL.PL 3PL.IO-more-old-ABS young-ADV die-PST
Kare brothers   oldest.of.them   as.young   he.died
'The eldest of the Kares brothers died young.'

(28) **ze-š-ze-šəpχʷə-xe-m**   *a-š'ə-š'-ew*   *psaw-ew*
REC.PR-brother-REC.PR-sister-PL-OBL 3PL.IO-LOC-be.part-ADV live-ADV
brothers.and.sisters.   being.part.of.them   alive
'Of all (our) brothers and sisters only Sveta and I are alive.'

(29) *he-m-re*   *č'etəw-m-re*   *ze.re-mə-ṣe-xe*
dog-OBL-COORD cat-OBL-COORD REC.ERG-NEG-know-PL
dog   cat   they.do.not.know.each.other
zə-χʷ-č'e,   **ze-pəj-xe-m**   *fe-d-ew*
REL.TMP-happen-INS REC.PR-enemy-PL-OBL BEN-correspond-ADV
if.it.happens   enemies   like
ze-fə-š'ə-tə-x…
REC.IO-BEN-LOC-stand-PL
they.stand.to.each.other
'If a dog and a cat have not been acquainted, they behave like enemies with respect to each other.'

The reciprocal prefix appears both with nouns occurring in the inalienable possessive construction, such as 'brother', and with alienable nouns which normally require the possessive prefix in the possessive construction, such as 'enemy'. Curious-

ly, with the latter, the construction lacks the possessive prefix (see Section 1) expected if this were just a reciprocal cross-reference prefix replacing the possessor (29). The suffixal part of the construction is not fixed: it can contain either a combination of the plural suffix with a case marker, as in (28)–(29), or just the oblique plural suffix, as in (27). On the other hand, the construction is impossible in the absence of plural morphology, so forms like *ze-šə-m* [REC.IO-brother-OBL] are infelicitous. The reciprocal prefix is absent when the set is not closed, i.e. there exists an individual who satisfies the expressed relation but is not included in the reference of the nominal:

(30) *a-š' fe-de* **šəpχʷə-me** *a-r-a-ʔʷ-a?*
 that-OBL BEN-correspond sister-OBL.PL 3PL.IO-DAT-3PL.ERG-say-Q
 it like sisters they.say.it.to.them
 'Is that a thing to say to one's sisters?'

Unlike in languages with more canonical dyadic constructions, non-symmetric relational nouns do not serve as bases for this pattern.

2.3.5 Associative plural and other kinds of non-homogeneous plurality

The associative plural in West Circassian is expressed by the suffix *-txe* / *-dəxe* (Kumakhov 1971: 25; Ilyina 2011; Yaroslavtseva 2017). This suffix may attach to names to indicate groups of people – a family, a company of friends, etc. – associated with a certain person (focal referent in terms of Daniel & Moravcsik 2005):

(31) *mwa.rə* **meleχʷe hasane-txe-r** *gʷəbʁʷe-m jə-ha-ʁe-x* ...
 here.is Melekh Hasan-APL-ABS field-OBL LOC-go.in-PST-PL
 here.is Melekh Hasan(pl) field they.came.out.there
 'Here are Melekh Hasan and his family, (who) have come out to the field ...'

(32) **məhamwede-txe-r** *jež'aṗe-m š'ə-ʔe-x*
 Mohammed-APL-ABS school-OBL LOC-be-PL
 Mohammed(pl) school they.are.there
 'Mohammed and his friends are in school.'

(33) **gʷeš'paq-txe-me** *s-a-ʔʷə-č'a-ʁ*
 Goshpak-APL-OBL.PL 1SG.ABS-3PL.IO-LOC-meet-PST
 Goshpak(pl) I.met.with.them
 'I met Goshpak and her family.' [E]

The last example shows that the associative plural suffix may combine with the oblique plural marker.

Less clear is the ability of the associative plural suffix to modify other kinds of nominals. According to Yaroslavtseva (2017), who studied the associative plural in the Bzhedug dialect, the suffix appears with demonstrative pronouns in the series which usually means 'other' (34), pet names (35) and some kinship terms (*s-šəpχʷə-txe-r* 1SG.PR-sister-APL-ABS 'my sister and those who are with her'). This does not seem possible in Standard West Circassian.[6]

(34) **mə-d-re-txe-r**   *wəna-č̣'e-m*   *kʷeš'ə-ž'ə-ʁe-x*
      this-OTHER-ADJ-APL-ABS house-new-OBL move-RE-PST-PL
      other(pl)   new.house   they.moved
      'This one and those who are with him/her moved to a new house.' [E]
      (Yaroslavtseva 2017: 9)

(35) *š'agʷə-r* *čəʔe-ŝ,*   **ž'eke-txe-r**   *čəʔe me-stə*
      yard-ABS cold-CS   Jacka-APL-ABS cold DYN-burn
      yard   it.is.cold Jacka(pl)   cold (they).feel
      'It is cold in the yard, so Jacka (the dog) and the other pets feel cold.' [E]
      (Yaroslavtseva 2017: 10)

Normally the focal referent should be a member of the set referred to by the associative plural form. In the following example the appearance of the associative plural form is considered inappropriate because the focal referent is explicitly separated from the set which the form denotes:

(36) *asłan ŝʷefə-m* *ʔʷef* *š'-j-e-ŝe.*   **#asłan-txe-r** *mezə-m*
      Aslan plain-OBL work LOC-3SG.ERG-DYN-do Aslan-APL-ABS wood-OBL
      Aslan plain   work he.does.it.there   Aslan(pl)   wood
      zə-š'-a-ʁe-psefə-x
      RFL.ABS-LOC-3PL.ERG-CAUS-calm-PL
      they.rest.there
      Expected: 'Aslan works in the field. Aslan's family rests in the woods.' [E]

Quite commonly, associative plural nominals refer to the home of the focal referent. This meaning presumably results from the extension of the "family interpretation". In these contexts, the associative plural form does not literally imply reference to a group involving the focal referent (37), does not trigger plural indexing and does not require the appearance of a case marker (38):

---

6 Interestingly, the restrictions which West Circassian puts on its associative plurals are in sharp contrast to the parallel construction in the closely related Kabardian. The latter uses a different associative plural marker which combines with a broad range of nominals and allows even inanimate and generic focal referents (Bagirokova et al. 2020).

(37) ***mwəstafe-txe-m*** *djə tə-š'ə-ʔa-ʁ, aw jež' jə-sə-ʁ-ep*
Mustafa-APL-OBL at 1PL.ABS-LOC-be-PST but self LOC-sit-PST-NEG
Mustafa(pl)   at we.were.there   but himself he.did.not.sit.there
'We were at Mustafa's, but he himself was not there.' [E]

(38) ***qambweljet-txe*** *cəʁʷa-be qə-š'-a-ča.he* /
Kambolet-APL mouse-many DIR-LOC-3PL.ERG-run.around
Kambolet(pl) many.mice they.run.around.it
*q-a-š'-a-ča.he
DIR-3PL.IO-LOC-3PL.ERG-run.around
they.run.around.them
'A lot of mice run around Kambolet's home.' [E]

The sentence in (37) also shows that a pluralized noun with the associative plural suffix can act as the antecedent for anaphoric reference later in the sentence.

According to Ilyina (2011: 210), the associative plural suffix may be added to toponyms, from which nouns can be derived which refer to the place's inhabitants (*pš'əč'ewə-dəxe-r* Pshicho-APL-ABS 'the inhabitants of Pshicho'). Peripherally, it also expresses additive plurality – e.g., with *ate* 'father':[7]

(39) ***ŝʷ-jate-txe-r*** *qə-zə-ḳʷe-č'e mə*
2PL.PR-POSS+father-APL-ABS DIR-REL.TMP-go-INS this
your.father(pl) if.(they.)come this
wəne-m š'ə-d-ʁe-heč'e-š'tə-x
room-OBL LOC-1PL.ERG-CAUS-guest-FUT-PL
room we.will.receive.them.there
'If your (plural addressees) fathers arrive, we will receive them in this room.' [E]

Besides *-txe*, West Circassian may employ various reduplication models for the expression of non-homogeneous plurality (usually with the semantics of similative plural; cf. Daniel & Moravcsik 2005): cf. *š'emset~ẑʷamset* 'Shamset and/or people like her' (derived from a personal name), *č'ele~ʁʷale* 'youth' (derived from *č'ele-* 'boy'); cf. Abregov 2000, Lander 2016. Some of these derivations behave as collective nouns which are normally indexed by plural markers (see Section 4.1). Nonetheless, the context may profile an indeterminate individual rather than plurality, in which case a reduplicated nominal may refer to a singular individual (40). At the same time, we find that collective nouns derived by such reduplication may occasionally take additive plural morphology, as in (48) below (cf. also (25) above for similar forms of underived collective nouns).

---

7 Yaroslavtseva (2007) reports that *-txe* with demonstratives is also sometimes translated as additive plural in examples like (34).

(40) **ʔavtobus~ẑʷavtobus-gʷere** qe-re-kʷ!
    bus~SIM-some   DIR-WIMP-go
    some.bus-like   let.it.come
    'Let some sort of bus (or bus-like vehicle) come!' [E]

2.4 Verbal Number

The closest West Circassian comes to expressing verbal number is in some reciprocal constructions. The prefix *ze-* (see Table 1) has become frozen on certain predicative words or in combination with certain preverbs, meaning that the event itself implies multiple participants, which are backgrounded and cannot be expresed (see Letuchiy 2007 for details). For example, in (41) the reciprocal prefix emphasizes the multiplicity of pieces resulting from breaking the branch:

(41) qʷətame-r   **ze-**p-jə-čʼəčʼə-ʁ
    branch-ABS   REC.IO-LOC-3SG.ERG-break-PST
    branch   s/he.broke.it
    'S/he broke the branch (into parts).' [E]
    (Letuchiy 2007: 794)

3 Cross-reference and the Syntax of Number

3.1 Cross-reference proper

As stated in Section 1, in West Circassian arguments are normally cross-referenced on their heads. (42) demonstrates the patterns which are closest to canonical agreement as defined by Corbett (2006). There is an unmarked (singular) NP referring to an agent which corresponds to a singular ergative cross-reference prefix *jə-* on the predicate, and a plural indirect object NP corresponding to the 3rd person plural indirect object cross-reference prefix *a-*:

(42) a-šʼ   adəge-ṣenəʁeležʼ-nəbžʼəčʼe-xe-m ʔʷefə-šxwe
    that-OBL Circassian-scholar-young-PL-OBL work-AUG
    he   young.Circassian.scholars   big.work
    a-d-jə-ṣa-ʁ
    3PL.IO-COM-3SG.ERG-do-PST
    he.did.it.with.them
    'He undertook a lot of work with young Circassian scholars.'

Importantly, cross-reference in West Circassian has a consistently semantic basis and number (as well as person) is marked without regard to any formal features of

external NPs. First of all, such NPs are often absent, and it is unclear why they should be postulated. Thus, example (42) above would remain grammatical without any NPs, in which case the arguments of the predicate get a pronominal interpretation 'he did it with them'. More relevantly to the topic of this chapter, when the relevant NP is present, its number value may differ from the number of the corresponding cross-reference prefix. Apart from the examples where there is a quantifier implying plurality as well as examples where the number of a count NP remains unmarked (to be discussed later in Section 2.3.2), there is the following evidence.

First, an unmarked (for number) group NP, i.e. a count nominal which refers to a set of individuals, may be indexed not only by a singular but also by a plural cross-reference prefix (43); cf. Kumakhov (1971: 14–20). These two options presumably depend on whether the group is considered as a multiplicity of individuals or as a single atomic entity. This is supported by the fact that when a non-plural group noun must get an atomic interpretation because of the context, it requires singular indexing (44).

(43) **kom'iss'ije-m**   s-**a**-tje-ḳʷa-ʁ /   sə-**Ø**-tje-ḳʷa-ʁ
     committee-OBL 1SG.ABS-3PL.IO-LOC-go-PST 1SG.ABS-3SG.IO-LOC-go-PST
     committee   I.defeated.them   I.defeated.it
     'I defeated the committee.' [E]

(44) mə   komjəssjəje-m   jež'.jež'ə.r-ew zə-ze-x-**jə**-š'a-ʁ /
     this committee-OBL self-ADV   RFL.ABS-REC.IO-LOC-3SG.ERG-organize-PST
     this committee   on.its.own   it.organized.itself
     ???zə-ze-x-**a**-š'a-ʁ
     they.organized.themselves
     RFL.ABS-REC.IO-LOC-3PL.ERG-organize-PST
     'This committee organized itself.' [E]

Kumakhov (1971: 15–16) reports that when *č'ǝle* 'village' is unmarked for number but the corresponding cross-reference is plural, it may receive the interpretation 'inhabitants of the village' as well as some other interpretations: in fact, the construction where *č'ǝle* is indexed by a plural prefix sometimes is not even interpreted with respect to any village but is understood as referring to people in general. The following examples can be uttered even in a context where no village is meant.

(45) səd-a,   **č'ǝle-m**   ja-mə-ceqe-ʁe   hašχʷəreʔʷ te
     what-Q village-OBL 3PL.IO+DAT-NEG-bite-PST rabid   we
     what   village   which.did.not.bite.them rabid   we
     qə-t-e-ceqa-ʁ-a?
     DIR-1PL.IO-DAT-bite-PST-Q
     it.bit.us
     'What? Did a rabid animal, which did not bite any other people (lit., village), bite us?'[E]

(46) **č'əle-m**   **ja**-sabjəj-xe-r   **ja**-wəne-xe-m
     village-OBL 3PL.IO+POSS-child-PL-ABS 3PL.IO+POSS-house-PL-OBL
     village   their.children   their.houses
     a-rə-sə-x
     3PL.IO-LOC-sit-PL
     they.sat.in.them
     'People's (lit., The village's) children stay at home.' [E]

Second, unmarked collective nouns like 'police' and 'youth' may trigger plural in-
dexing, either optionally (47a) or obligatorily (48). At least for some speakers, col-
lective nouns themselves may take plural morphology in these contexts (cf. (24);
although sometimes they are not likely to take the plural suffix *-xe* (47b) and allow
only the oblique plural marker):

(47) a. **pol'icəje-m** *bzeǯ'aŝe-xe-r*   q-**a**-wəbətə-ʁ /   q-**ə**-wəbətə-ʁ
        police-OBL   *criminal*-PL-ABS DIR-3PL.ERG-catch-PST DIR-3SG.ERG-catch-PST
        police   criminals   they.caught.(them)   it.caught.(them)

     b. **pol'icəje-me** / ??**pol'icəje-xe-m** *bzeǯ'aŝe-xe-r*   q-**a**-wəbətə-ʁ
        police-OBL.PL   police-PL-OBL   criminal-PL-ABS DIR-3PL.ERG-catch-PST
        police(pl)   police(pl)   criminals   they.caught.(them)
        'The police caught the criminals.' [E]

(48) *a-xe-r*   **t-jə-č'ele~ʁʷale(-xe)-m**   **ja**-d-ʁa-ŝe-xe-me /
     that-PL-ABS 1PL.IO-POSS-youth-PL-OBL 3PL.IO+DAT-1PL.ERG-CAUS-do-PL-COND
     they   our.youth   if.we.teach.them.to.them
     *****Ø**-je-d-ʁa-ŝe-xe-me   t-ŝʷe-jəʁʷ-ew   a-wəž
     3SG.IO-DAT-1PL.ERG-CAUS-do-PL-COND 1PL.IO-MAL-wish-ADV 3PL.IO-after
     if.we.teach.them.to.it   it.is.our.wish   after.them
     t-jə-t
     1PL.ABS-LOC-stand
     we.stand.there
     'We are busy (lit., standing after) trying to teach these things to our youth.'

Third, as was first noticed by Sergei Minor and further discussed by Arkadiev &
Lander (2013), distributive NPs containing the quantifier 'every' may correspond
either to singular (49) or plural (50) indexing, with no visible semantic difference.[8]

---

8 Another distributive quantifier discussed by Arkadiev and Lander (2013) is *qes* 'every' (used with
time periods). We have not come across any natural example where a phrase containing this quanti-
fier would correspond to an overt index.

(49) *šxən-xe-r*   **wənaʁʷe-pepč**   *Ø-š'-a-ṣ̂ə-š'tə-ʁe-x*
meal-PL-ABS   family-every   3SG.IO-LOC-3PL.ERG-do-AUX-PST-PL
meals   every.family   they.make.them.in.it
'Meals were cooked in every household.' [E]

(50) *ǯ'a-š'*   *fe-de*   *gʷəpč'e-xe-r*   **rajon-pepč**
that-OBL   BEN-correspond   centre-PL-ABS   district-every
that   like   centres   every.district
a-*š'-a-ʁe.psə-n-ew*   *a-ʁe-nafe*
3PL.IO-LOC-3PL.ERG-create-MOD-ADV   3PL.ERG-CAUS-clear
in.order.for.them.to.create.(them).in.them   they.decided.it
'They decide to create centres like that in every district.'

If indexing here is interpreted as agreement, it would violate Tatevosov's (2002: 80) generalization that NPs containing distributive quantifiers should be associated with singular agreement (unless they contain a plural noun). To comply with Tatevosov's arguments, this construction must be understood as evidence that plural indexing in such examples is not agreement at all.

Fourth, we find a peripheral inclusory construction, where the 1st person plural indexing is linked to coordinate NPs referring to the subset of the relevant argument. In the following example we find the coordinate construction 'Makhmoud and Kaplan' associated with the 1st person plural indexing and clearly forming a subset of the absolutive argument of the final predicate:

(51) *səd-a,*   *w-jə-zakʷ-ew*   *wə-qə-š'ə-ŝʷa-ʁ-a?*
what-Q   2SG.IO-POSS-alone-ADV   2SG.ABS-DIR-LOC-dance-PST-Q
what   you.alone   you.danced.there
s-jə-zeqʷe-n-a?   **maħmwəd-jə,**   **qepλan-jə**   **(*te)**
1SG.IO-POSS-alone-MOD-Q   Makhmoud-ADD   Kaplan-ADD   we
me.alone   Makhmoud   Kaplan   we
tə-*qə-š'ə-ŝʷa-ʁ*
1PL.ABS-DIR-LOC-dance-PST
we.danced.there
'What, did you dance by yourself? – What do you mean, by myself? I (lit., we) danced with Mahmoud and Kaplan.' [E]

In addition to being limited to the 1st person plural indexing, this construction is also subject to the following constraints. First, at least two other grammatical participants must be included besides the first person, joined by a coordination marker. Second, the dedicated coordination marker *-re* cannot be used, rather, the general additive marker *-jə* is required. Finally, the (singular or plural) first person may not be expressed with a free pronoun.

Whenever the number is irrelevant (e.g., in most generic contexts), indexing corresponds to whether or not the corresponding nominal is marked as plural. Thus, deficient nouns are undefined for number, do not take plural morphology and do not allow plural indexing:

(52) **francuz** *neʔʷase*   *sə-**Ø**-fe-χʷə-n-ew /*
     French   acquaintance 1SG.ABS-3SG.IO-BEN-happen-MOD-ADV
     French   acquaintance for.me.to.become.to.him/her
     s-a**-fe-χʷə-n-ew*   *sə-fa-j*
     1SG.ABS-3PL.IO-BEN-happen-MOD-ADV 1SG.ABS-BEN-want
     for.me.to.become.to.them   I.want.it
     'I want to meet French people.' [E]

Furthermore, 3rd person plural indexing may be used where the argument is impersonal. In such cases, the appearance of a (pronominal) NP is impossible:

(53) *dejə-ç'e* *č'ale-m*   *ja-gʷəʁʷ*   (#*a-xe-me*)
     bad-INS   village-OBL 3PL.IO+POSS-mention   that-PL-OBL.PL
     bad   village   their.mention   they
     *š'-**a**-ŝə-ʁ-ep*
     LOC-3PL.ERG-do-PST-NEG
     they.did.not.do.there
     'In the village nothing bad was said about them.'

3.2 Plural suffix

Yet another plural marker which appears on the predicate is the suffix *-xe*, probably the same morpheme that marks plurality in NPs. Typically it is used on predicates with 3rd person plural absolutives (recall that the 3rd person absolutive is not indexed by prefixes), but even here it is optional (54). Predicate nominals do not differ from other predicates in this respect, see Section 2.3.2. The plural suffix can also appear when the absolutive is expressed by the reflexive prefix controlled by some other plural argument, as in (55), see also Rogava & Kerasheva (1966: 167).

(54) *čəgʷ-ŝabe-m*   *š'ə-š'*   ***a-xe-r***   *me-χʷə-ž'ə(-**x**)*
     earth-soft-OBL LOC-be.part that-PL-ABS   DYN-happen-RE-PL
     soft.earth   part.of.it   they   they.became
     'They become part of the soft earth.'

(55) *newəžə-m* *prakt'ič'eske-ze.neqʷeqʷə-xe-m* *a*
     later-OBL   practical-contest-PL-OBL   that
     later   practical.contests   that

qə-x-a-χə-ʁe-xe-m   ***z**-a-š'-a-wəšetə-š'tə-**x***
DIR-LOC-3PL.ERG-take.out-PST-PL-OBL   RFL.ABS-3PL.IO-LOC-3PL.ERG-test-FUT-PL
ones.whom.they.chose   they.will.test.themselves.in.them
'Later those who were chosen will test themselves in contests of practical skills. '

Occasionally, the use of *-xe* on predicates is also extended to forms containing 1st and 2nd person plural absolutive cross-reference:

(56) *ʁəš'-xe-r,*   *xeʷkəẑ-xe-r,*   *sṭaŝʷ-xe-r,*   *matχəʁe-xe-r,*
Gish-PL-ABS   Khakuz-PL-ABS   Stash-PL-ABS   Mamkheg-PL-ABS
Gishes   Khakuzhes   Stashes   Mamkhegs
fe-ŝhaf-xe-r-jə   *adəgjejə-m*   ***tə**-qe-ḳʷa-ʁe-**x***
BEN-other-PL-ABS-ADD   Adyghea-OBL   1PL.ABS-DIR-go-PST-PL
and.others.than.them   to.Adyghea   we.came
'We, Gishs, Khakyzhs, Stashes, Mamkhegs and others, came to Adyghea.'

The rules formulated above for 3rd person plural cross-reference seem to work for plural suffixation as well (in the sense that these rules reflect the referential properties of the relevant arguments as opposed to the formal properties of the corresponding NPs). The following examples illustrate the appearance of the plural suffix for arguments that correspond to group nouns (57) and nominals containing a distributive quantifier (58):

(57) ***zə**   **ḳʷəpə-r***   *wəneʁʷe-baj-dede-xe-m*   *q-a-re-č'ə-**x***
one   group-ABS   family-rich-very-PL-OBL   DIR-3PL.IO-TRANS-go.out-PL
one group   very.rich.families   they.came.from.them
'One group is from very rich families.'

(58) ***ə-ʁe-wəneχʷə-ʁe-pepč**,*   *sə-č'e-wəpče-ze,*   *qema-pe-č'e*
3SG.ERG-CAUS-miserable-PST-every   1SG.ABS-LOC-ask-CNV   dagger-nose-INS
everyone.whom.he.made.miserable   while.asking   with.a.dagger
*qə-xe-z-ʁe-pəž'e-š'tə-**x***
DIR-LOC-1SG.ERG-CAUS-stab-FUT-PL
I.will.allow.them.to.stab.him
'While making inquiries, I will allow everyone whom he made miserable to stab him with a dagger.'

Like 3rd person plural prefixes, the plural suffix appears when the absolutive is impersonal (and normally is not expressed with an independent NP), as in (59). In the impersonal construction the presence of *-xe* is obligatory.

(59) *qə-w-e-ǯ'e-x*
DIR-2SG.IO-DYN-call-PL
they.call.you
'You are being called.' [E]

3.3 NP-internal marking

In general, there is no NP-internal agreement in West Circassian. The only possible exception is found in relative clause constructions.[9] The plural suffix *-xe* discussed in Section 3.2 may appear on predicates of relative clauses, in which case it marks the plurality of the absolute argument (60) or the argument being relativized (61). When compared to these examples, (62) demonstrates that the plural suffix is optional in both cases: here the relativized argument is the absolute, but the plural is not marked in 'what they see' (that it is plural is known from the plural suffix on the predicate of the matrix clause 'they retold them').

(60) *marə*   ***zə-de-ḳʷa-ʁe-xe-r***
     here.is REL.IO-LOC-go-PST-PL-ABS
     here.is where.they.went
     'Here is (the place) where they went.' [E]

(61) ***s-je-zə-ʁe-ǯ'a-ʁe-xe-m***   *tha-ŝʷ-je-ʁe-psew*
     1SG.ABS-DAT-REL.ERG-CAUS-read-PST-PL-OBL God-2PL.IO-DAT-CAUS-live
     ones.who.taught.me   thank.you
     ja-s-ʔʷe-ž'ə   *s-ŝʷe-jəʁʷ*
     3PL.IO+DAT-1SG.ERG-say-RE   1SG.IO-MAL-wish
     I.tell.them   my.wish
     'I want to say thank you to those who taught me.'

(62) *pšəse-xe-r-jə*   *ze-x-a-λha-š'tə-ʁe-x,*   ***a-λeʁʷə-re-r-jə***
     tale-PL-ABS-ADD   REC.IO-LOC-3PL.ERG-put-AUX-PST-PL 3PL.ERG-see-DYN-ABS-ADD
     tales   they.composed.them   what.they.see
     q-a-ʔʷete-ž'ə-š'tə-ʁe-x
     DIR-3PL.ERG-tell-RE-AUX-PST-PL
     they.retold.them
     'They wove tales, retold what they had seen.'
     (Lander 2012: 234–235)

---

9 This section is largely based on Lander 2012 (chapter 3).

The examples just given represent the "free relative" construction, i.e. a construction lacking a semantic head (a nominal semantically modified by the relative clause). When there is a nominal serving as the semantic head of the construction, the picture becomes more complex.

The semantic head of the relative clause construction in West Circassian may be internal or external. In the internally-headed relative clause construction, the semantic head takes the adverbial suffix and apparently belongs to the relative clause (more often than not on its left periphery). If the referent of the whole construction is plural, the internal head can but need not be marked for plural (63). It can also be represented by a nominal containing a numeral (64), i.e. an element that does not normally combine with plural marking in the NP (Section 2.3.2). In such cases, however, the relative clause predicate, which here simultaneously functions as the syntactic head of the NP, is normally marked with the plural suffix.

(63) ***wəpče[-x-e]w***   ***š'ə-tə-ʁe-xe-m***   *zeč'e-m-jə*   *a-š'*   *ǯ'ewap-xe-r*
    question-PL-ADV   LOC-stand-PST-PL-OBL   all-OBL-ADD   that-OBL   answer-PL-ABS
    question(s)   one(s).which.stood   all   he   answers
    qə-ze.r-a-r-jə-tə-ž'ə-ʁe-r   *q-ə-ʔʷa-ʁ*
    DIR-REL.FACT-3PL.IO-DAT-3SG.ERG-give-RE-PST-ABS   DIR-3SG.ERG-say-PST
    what.he.gave.them   he.said.it
    'He said that he gave them answers to all the questions that (they) had.'

(64) ***ʁʷep-jə-ṭʷ-ew***   ***jə-ʔe-xe-r***   *šχʷanṭe-x*
    sleeve-LNK-two-ADV   POSS-be-PL-ABS   blue-PL
    two.sleeves   what.he.had   they.are.blue
    'The two sleeves it has are dark blue.'

In the externally-headed construction, the semantic head follows the relative clause and displays usual head properties including marking of number (Lander 2012); cf. (65). Yet, marking of the plurality of the relativized argument may occasionally be doubled on the predicate of the relative clause. Still, plural marking only on the predicate but not on the external head is very infrequent (66).

(65) ***a-xe-m***   ***a-čəpe***   ***q-jə-he-re***   ***č'ele-çaḳʷə-xe-r*** *jəč'erač'-ew*
    that-PL-OBL   3PL.PR-place   DIR-LOC-go.in-DYN   boy-small-PL-ABS   anew-ADV
    they   their.place   who.enter   children   anew
    q-je-w-e-ʁe-ž'e-ž'ə-ŝ   *w-e-ʁe.se-ž'ə-x*
    DIR-DAT-2SG.ERG-DYN-CAUS-depart-RE-CS   2SG.ERG-DYN-educate-RE-PL
    you.start.it   you.educate.them
    'You educate the children who replace them, starting this from the very beginning.'

(66) *ʎesə*   *zə-zə-ṣ̂ə-ʁe-xe*   *zeweʎ-me* /   ???*zeweʎə-m*
     pedestrian RFL.ABS-REL.ERG-do-PST-PL warrior-OBL.PL warrior-OBL
     pedestrian who.made.themselves   warriors   warrior
     ze-x-a-x-ew   *q-ə-ʔʷa-ʁ*
     REC.IO-LOC-3PL.ERG-hear-ADV DIR-3SG.ERG-say-PST
     they.hear.it   he.said.it
     'He said it so that the warriors who made themselves dismount heard it.'

Such doubling may be described as optional agreement of the predicate of the relative clause with the external head. Yet it is more likely that the plural marking in this case results not from agreement but from the plurality of the relativized argument, as is the case with free relatives (cf. (61)).

4 Semantics and Discourse

4.1 Number interpretation and referential properties

In this section we describe the behavior of the category of number for various nominals which differ in their referential properties. We are mostly interested in how the interpretation of number forms depends on such features as specificity, definiteness, genericity, etc. In this section we focus on count nouns.

     Indefinite uses. As shown in Section 2.3.2, being non-specific, deficient NPs are undefined for number. Still, not all non-specific NPs are deficient, and a non-specific NP can be explicitly marked as plural, as in (67):

(67) *qə-z-de-ʔepəʔe-š't ə-xe-m*   *s-a-ʎe-χʷə.*   *xet-m-jə*
     DIR-REL.IO-COM-help-FUT-PL-OBL 1SG.ABS-3PL.IO-LOC-search who-COND-ADD
     ones.who.will.help.me   I.look.for.them   anyone
     χʷə-š't
     happen-FUT
     s/he.will.do
     'I'm looking for someone to help me. Anyone will do.' [E]

A specific indefinite NP is normally defined with respect to the category of number. For example, when a phrase unmarked for number can serve as an antecedent for anaphora (such an ability being a prerequisite for specificity), it should refer to a singular individual. A sentence like (68a) can be continued by (68b) but not by (68c):

(68) a. **qə-z-de-ʔepəʔe-š'tə-m** sə-λe-χʷə …
DIR-REL.IO-COM-help-FUT-OBL 1SG.ABS-LOC-search
someone.who.will.help.me I.look.for.him/her
'I am looking for someone who will help.' [E]

b. … **a-r** zə-gʷere-m ča-ʁe
that-ABS one-some-OBL run-PST
s/he somewhere ran
'S/he has run somewhere.' [E]

c. … #**a-xe-r** zə-gʷere-m ča-ʁe-x.
that-PL-ABS one-some-OBL run-PST-PL
they somewhere ran
'They have run somewhere.' [E]

In general, it seems that whenever a salient discourse referent is introduced, its number is explicitly marked. For example, it is impossible to leave the number unmarked in presentational constructions intended to introduce protagonists. Consequently, the subject NP in the first sentence of (69a) cannot serve as the antecedent of a plural pronoun (69c), but only of a number-unmarked pronoun (69b) (here interpreted as singular):

(69) a. ze-gʷere-m psewə-š'tə-ʁe **ʁʷəč'e-ʔaze**…
once-some-OBL live-AUX-PST blacksmith-skilled
once s/he.lived skilled.blacksmith
'Once there lived a skilled blacksmith.' [E]

b. … **a-r** zeč'e-m-jə ze-λ-a-ŝe-š'tə-ʁe
that-ABS all-OBL-ADD REC.IO-LOC-3PL.ERG-know-AUX-PST
s/he all they.knew.him/her
'Everyone knew him.' [E]

c. … #**a-xe-r** zeč'e-m-jə ze-λ-a-ŝe-š'tə-ʁe-x
that-PL-ABS all-OBL-ADD REC.IO-LOC-3PL.ERG-know-AUX-PST-PL
they all they.knew.him/her
'Everyone knew them.' [E]

Kind-referring, generic, and generic-like contexts. Following Krifka et al. (1995), we distinguish between reference to kinds and reference to objects. The difference is related to the selectional restrictions of predicates (some predicates cannot apply to individual instantiations of a kind) and, in some languages, may restrict the range of available constructions and forms.

Kind-referring NPs other than referring to the plurality of species (see below) prefer the unmarked form, though the plural form is often allowed as well. Cf. (70)–(71), which contain a predicate that takes a kind-referring argument:

(70) *çəf-xe-m*   *beʂaʁ-ew*   **š'ereχə-r** / ??**š'ereχə-xe-r**
     person-PL-OBL long.ago-ADV wheel-ABS   wheel-PL-ABS
     people   long.ago   wheel   wheels
     q-a-wəgʷəpšəsə-ʁ
     DIR-3PL.ERG-devise-PST
     they.invented.it
     'People invented the wheel long ago.' [E]

(71) a. **ha-r** /   **ha-xe-r**   *təʁʷəẑə-m* *qə-tje-čʼə-ʁ*
        dog-ABS dog-PL-ABS wolf-OBL   DIR-LOC-go.out-PST
        dog   dogs   wolf   s/he.went.out.of.it

     b. **ha-r** /   **ha-xe-r**   *təʁʷəẑ-xe-m* *q-a-tje-čʼə-ʁ*
        dog-ABS dog-PL-ABS wolf-PL-OBL   DIR-3PL.IO-LOC-go.out-PST
        dog   dogs   wolves   s/he.went.out.of.them
        'The dog is descended from the wolf.' [E]

Kind-referring plurals can also get taxonomic readings, i.e. refer to a set of species (72), similarly to sort reading of the plurals with mass nouns (see Section 2.3.4).

(72) **šə-xe-r**   *ze.r-a-ʁe-fjede-re-m-čʼe*
     horse-PL-ABS REL.MNR-3PL.ERG-CAUS-use-DYN-OBL-INS
     horses   by.the.way.they.use.(them)
     ze-tje-čʼə-x
     REC.IO-LOC-go.out-PL
     they.go.out.on.each.other
     'Horses are differentiated by how they are used.' [E]

Kind-referring NPs may (72) – but need not (70)–(71) – occur in generic statements. On the other hand, generic statements (73) and statements containing adverbs of quantification like 'always' (74) may contain object-referring NPs. Thus, in (73) the generic statement suggests that dogs by default eat meat, but they do so individually (cf. (70), where it is not true that each individual wheel has been invented by humans). In the standard register, such NPs are normally case-marked but allow both the presence and the absence of plural morphology. Occasionally complex deficient phrases are considered possible in this context as well (75).

(73) a. **ha-xe-m**   *lə*   *a-šxə,*   **šə-xe-m**   *a-šxə-r-ep*
        dog-PL-OBL meat 3PL.ERG-eat horse-PL-OBL 3PL.ERG-eat-DYN-NEG
        dogs   meat they.eat.it   horses   they.do.not.eat.it

     b. **ha-m**   *lə*   *j-e-šxə,*   **šə-m**   *ə-šxə-r-ep*
        dog-OBL meat 3SG.ERG-DYN-eat horse-OBL 3SG.ERG-eat-DYN-NEG
        dog   meat s/he.eats.it   horse   s/he.do.not.eat.it

c. ??**ha** *lə*   *j-e-šxə,*   **šə**   *ə-šxə-r-ep*
   dog   meat 3SG.ERG-DYN-eat horse 3SG.ERG-eat-DYN-NEG
   dog   meat s/he.eats.it   horse s/he.do.not.eat.it
   'Dogs eat meat, but horses don't.' [E]

(74) a. **adəge-xe-r**   *jeʁaŝ-jə*   *čəgʷə-m*   *d-e-laž'e-x*
     Circassian-PL-ABS always-ADD ground-OBL LOC-DYN-work-PL
     Circassians   always   ground   they.work.in.it

   b. **adəge-r**   *jeʁaŝ-jə*   *čəgʷə-m*   *d-e-laž'e* [E]
     Circassian-ABS always-ADD ground-OBL LOC-DYN-work-PL
     Circassian   always   ground   s/he.work.in.it

   c. *****adəge**   *jeʁaŝ-jə*   *čəgʷə-m*   *d-e-laž'e* [E]
     Circassian always-ADD ground-OBL LOC-DYN-work-PL
     Circassian always   ground   s/he.work.in.it
     'Circassians always practice agriculture.'

(75) **ʔaze-deʁʷ**   *jeʁaŝe-m*   *j-e-ʁe-χʷə-ž'ə*
   doctor-good   always-OBL 3SG.ERG-DYN-CAUS-happen-RE
   good.doctor   always   s/he.cure.him/her
   '(S)he is always treated by good doctors.' [E]

As compared to specific referents, in such examples the relevant NPs are undefined for number because speakers do not commit themselves to thinking of the number of referents. Still, the use of plural morphology implies that the set of objects relevant to the utterance is likely to include more than one element.

Interrogative uses. West Circassian has several pronouns which are primarily used as question words – like *xet* 'who' (used only for humans), *səd* 'what', *təde* 'where', *sədjəʁʷe* 'when', *thapš* 'how many'. In questions, these pronouns typically appear as predicates of a pseudocleft (76)–(77), but under some conditions – e.g., when the event is not presupposed (cf. (78), and see Sumbatova 2009) or when the questioned participant is deeply embedded – they may remain in situ and acquire the relevant case marking.

(76) **xet-a** *cace-č'e*   *tχʷə*   *zə-ʁe-ẑa-ʁe-r?*
   who-Q fork-INS   butter REL.ERG-CAUS-fry-PST-ABS
   who   with.a.fork butter one.who.fried.it
   'Who is the one who fried butter with a fork?'

(77) *mə*   *λenəqʷe-m-č'e*   **səd-a**   *ŝʷ-ŝa-ʁe-xe-r?*
   this direction-OBL-INS what-Q 2PL.ERG-do-PST-PL-ABS
   this in.direction   what   what(pl).you.did
   'What did you do in this direction?' (Lit., 'In this direction, what is those that you did?')

(78) *dwənajə-r qʷəte-ž'ə-ʁ-a ŝəw, **səd** qe-χʷə-ʁ-a?*
world-ABS break-RE-PST-Q PTCL what DIR-happen-PST-Q
world   it.broke   PTCL what it.happened
'Has the world ended or what? What/Anything happened?' [E]
(Sumbatova 2009: 601)

Question words are neutral for number. Yet if the plurality of the relevant argument is presupposed, the plural of the questioned argument can be expressed elsewhere, as in (77), where it is marked by the plural suffix on the headless relative. Interrogative pronouns may also receive plural marking themselves, both when functioning as predicates and when referring to arguments. This is illustrated for 'who' in (79), for 'what' in (80) and for 'how many' in (81), where examples (a) demonstrate the predicate position and other examples show case-marked uses:

(79) a. ***xetə-x-a*** *č'elejež'aḳʷe-xe-r a-nahə-b-ew*
who-PL-Q pupil-PL-ABS   3PL.IO-more-many-ADV
who(pl)   pupils   most.of.them
qə-če-zə-ṭʷəpš'ə-ʁe-xe-r?
DIR-LOC-REL.ERG-let.go-PST-PL-ABS
one.who.graduated
'Who graduated the most pupils?' (Lit., 'Who is that who graduated the most pupils?')

b. *çəf-λepq-ç'e   we   **xet-me**   w-a-š'ə-š'?*
person-kind-INS you.sg who-OBL.PL 2SG.ABS-3PL.IO-LOC-be.part
by.nation   you   who(pl)   you.are.part.of.them
'What is your nationality?' (Lit., 'Of whom (plural) are you a part of by your nation?')

c. ***xet*** / ***xet-xe-m*** *a-djə   wə-š'ə-ʔa-ʁ-a?*
who who-PL-OBL 3PL.IO-at 2SG.ABS-LOC-be-PST-Q
who who(pl)   at.them   you.were.there
'At whose house were you?' [E]

(80) a. *ʔʷef-x-ew   çəf-xe-m   ja-b-ʁa-ŝe-xe-re-r*
work-PL-ADV person-PL-OBL 3PL.IO+DAT-2SG.ERG-CAUS-do-PL-DYN-ABS
jobs   people   what.you.make.the.do
səd-x-a?
what-PL-Q
what(pl)
'What are the jobs that you make people do?'

b. ***səd-xe-m*** *a-pa.je* *mə ʔaλmeq-xe-r qe-p-šta-ʁ-a?*
what-PL-OBL 3PL.IO-for this bag-PL-ABS DIR-2SG.ERG-take-PST-Q
what(pl) for.them this bags you.took.them
'For what reasons have you taken these bags?' [E]

(81) a. *təʁeʁaz-ew,* *natrəf-ew xe-ŝʷ-ʎha-ʁe-xe-r* ***thapš-x-a?***
sunflower-ADV corn-ADV LOC-2PL.ERG-put-PST-PL-ABS how.many-PL-Q
sunflower corn what.you.planted how.many
'How many are the sunflowers and corn you planted?'

b. ***wəpče-thapš-me*** *ǯ'ewap ja-p-tə-ž'ə-ʁ-a?*
question-how.many-OBL.PL answer 3PL.IO+DAT-2SG.ERG-give-RE-PST-Q
how.many.questions answer you.gave.it.to.them
'How many questions did you give an answer to?'

At least when referring to time, pluralization of 'how many' can convey a sense of approximation:

(82) *hač'-ew* *tə-z-a-že-re-m* ***thapš-xe-m*** *a-djə*
guest-ADV 1PL.ABS-REL.IO-DAT-wait-DYN-OBL how.many-PL-OBL 3PL.IO-at
guest one.whom.we.wait.for how.many(pl) at.them
qe-ķʷe-n-ew *ə-ʔʷa-ʁ-a?*
DIR-go-MOD-ADV 3SG.ERG-say-PST-Q
s/he.will.come s/he.said.it
'The guest we are waiting for said he would arrive at approximately what time?' [E]

The pronouns *təde* 'where' and *sədjaʁʷe* 'when' cannot be pluralized unless they are constructed in parallel to a plural expression that the speaker asks to repeat:

(83) *a* *zeman-xe-m t-jə-č'əle-me* *ʁeŝeʁʷen-xe-r*
that time-PL-OBL 1PL.IO-POSS-village-OBL.PL wonderful-PL-ABS
that times our villages miracles
qə-š'ə-χʷə-š'tə-ʁe.
DIR-LOC-happen-AUX-PST
they.happened.there
sədjaʁʷe-xe-m *p-ʔʷa-ʁ-a?* ***təde-me*** *p-ʔʷa-ʁ-a?*
when-PL-OBL 2SG.ERG-say-PST-Q where-OBL.PL 2SG.ERG-say-PST-Q
when(pl) you.said.it where(pl) you.said.it
'In those days in our villages miracles happened. – When, you say? Where, you say?' [E]

Besides questions, question words can be used in some functions typical for indefinite pronouns (cf. Haspelmath 1997), e.g., in general concessive clauses (84) or in

free choice contexts (85). In the former, they cannot trigger plural indexing in the subordinate clause (and hence behave as deficient nouns), while in the latter they trigger the presupposition of plurality and normally require plural indexing:

(84) ***xet*** *ž'ane* (*a-)*fe-z-də-ʁe-m-jə,*   *zeč'e-r-jə*
    who dress 3PL.IO-BEN-1SG.ERG-sew-PST-COND-ADD all-ABS-ADD
    who dress I.sewed.it.for.him/her   all
    gʷəŝʷe-š'tə-ʁe
    happy-AUX-PST
    s/he.was.happy they.were.happy
    'No matter for whom I sewed a dress, they were all happy.' [E]

(85) ***xet-m-jə***   *(*a*-)*fe-z-də-ʁ*
    who-COND-ADD   3PL.IO-BEN-1SG.ERG-sew-PST
    whoever   I.sewed.it.for.them
    '(He) sewed for anyone.' [E]

Quantified NPs. With many quantifiers whose use implies plurality and which can trigger plural indexing, overt plural marking is possible but optional. The following sets of examples illustrate unmarked for number (a) and plural (b–c) NPs containing quantifiers *zawəle* 'several' (86), *pčaʁe* 'a number of', literally 'number' (87), *be* 'many' (88) and *pstew* 'all' (89) (see Nikolaeva 2012 for a detailed description of West Circassian quantifiers).

(86) a. ***mefe-zawəle-m*** *merzəχʷaj*   *qe-ʎeʁʷa-ʁ-ep*
        day-several-OBL Marzukhay DIR-be.seen-PST-NEG
        several.days   Marzukhay he.was.not.seen
        'Marzukhay did not show himself for a few days.'

   b. ***mefe-zawəle-xe-m*** *weš'x-ṭekʷ-xe-r*   *q-je-š'xə-š'tə-x*
        day-several-PL-OBL rain-little-PL-ABS DIR-DAT-rain-FUT-PL
        several.days   light.rains   they.will.rain
        'A light rain will fall for a few days.'

(87) a. ... ***ʔeč'əb-qeraləʁʷe-pčaʁe-m***   *ja-scene-xe-m*
        external-country-number-OBL 3PL.IO+POSS-stage-PL-OBL
        number.of.foreign.countries   their.stages
        rol'-ʁeŝeʁʷenə-be   *q-a-š'ə-p-ŝə-ʁ*
        role-wonderful-many   DIR-3PL.IO-LOC-2SG.ERG-do-PST
        many.wonderful.roles you.did.on.them
        'You played many interesting roles on the stages of a number of foreign countries.'

b. ***qeraləʁʷe-pčaʁe-xe-m*** *a-š'ə-ʔa-ʁ,*   *koncert-xe-r*
    country-number-PL-OBL 3PL.IO-LOC-be-PST concert-PL-ABS
    number.of.countries   she.was.in.them   concerts
    q-a-š'-jə-tə-ʁe-x
    DIR-3PL.IO-LOC-3SG.ERG-give-PST-PL
    she.gave.them.in.them
    'She went to (lit., was in) a number of countries and gave concerts there.'

c. ***qeraləʁʷe-pčaʁe-me***   *ekonom'ike-r* *a-š'ə-ze-tje-zə-ʁa-ʁ*
    country-number-OBL.PL economy-ABS 3PL.IO-LOC-REC.IO-LOC-fall-PST-PST
    number.of.countries   economy   it.fell.in.them
    'The economy collapsed in many countries.'

(88) a. ***çəfə-be-m***   *ŝʷə-č'e*   *zə-çe*   *r-a-ʔʷa-ʁe-r?*
    person-many-OBL good-INS   REL.PR-name DAT-3PL.ERG-say-PST-ABS
    many.people   with.kindness whose.name they.say.it.to.him
    'The one who many people spoke of with kindness?'

b. *a-xe-me*   *blaʁe-č'əž'e-xeʁegʷə-me*
    that-PL-OBL.PL near-far-country-OBL.PL
    they   near.far.countries
    ja-š'en-š'ef-çəfə-be-xe-r   *a-x-e-laž'e-x*
    3PL.IO+POSS-sell-buy-person-many-PL-ABS 3PL.IO-LOC-DYN-work-PL
    many.merchants.of.them   they.work.in.them
    'Many merchants from the near- and far-abroad participate in them.'

c. ***çəfə-be-me***   *zavodə-m* *ʔʷef* *š'-a-ŝe*
    person-many-OBL.PL factory-OBL work LOC-3PL.ERG-do
    many.people   factory   work they.do.it.there
    'Many people work at the factory.'

(89) a. ... *a-xe-me*   *ja-ʁʷəse-x-ew*   *the-m*
      that-PL-OBL.PL 3PL.IO+POSS-accompany-PL-ADV God-OBL
      they   with.them   God
      jə-çəf-ƛeṗe-pstewə-m   *a-ʔape*   *s-fe-ŝʷə-wəbətə-ž'*
      POSS-person-respected-all-OBL 3PL.PR-palm 1SG.IO-BEN-2PL.ERG-catch-RE
      all.respected.people.of.his   their.finger   you.catch.it.for.me
      'Salute (lit., catch the palm) [for me] all the saints (lit., the respected
      people of God) which are with them.'

b. *a* ***ŝʷeʁe-pstewə-me*** *a-paje*   *ƛepq-eposə-m*
    that merit-all-OBL.PL   3PL.IO-for nation-epic.poem-OBL
    that all.merits   for.them   national.epic.poem

> *dwəneje-kul'ture-m   čəṗe-ʁenefaʁe   š'-j-e-wəbətə*
> world-culture-OBL   place-definite   LOC-3SG.ERG-DYN-catch
> world.culture   definite.place   it.catches.there
> 'Thanks to all of these qualities the national epic poem has a visible place in world culture.'

Not all quantifiers behave this way, though. In particular, *zeč'e* 'all' (90) does not allow the omission of plural marking on nouns without changing the meaning to 'whole'. Still, it does not require plural marking if what is quantified is a free relative; cf. (91) where the plurality of the subject is nonetheless reflected by the plural suffix on the predicate.

(90) ***zeč'e ʔʷefaʁʷe(*-xe)-r*** *ze-ŝʷe-p-xə-n*   *p- λeč'ə-š't-ep*
    all   issue-PL-ABS   REC.IO-LOC-2SG.ERG-carry-MOD   2SG.ERG-can-FUT-NEG
    all   issues   you.will.solve   you.will.not.be.able
    'You will not be able to solve all the problems.'

(91) ***a-xe-m***   ***zeč'e   ja-ʔe(-xe)-r*** (...)   *qʷehaṗe-m š'ə-ʔe-x*
    that-PL-OBL   all   3PL.IO+POSS-be-PL-ABS   West-OBL   LOC-be-PL
    they   all   what.they.have   West   they.are.there
    'Everything they have (...) is in the West.'

When quantifiers are used without explicitly mentioning the quantified, the plural morphology usually becomes obligatory, probably because most quantifiers do not imply plurality (they can combine with mass nouns) per se. Cf. (92)–(93), where the plural suffix cannot be omitted; but see (94), where plural marking becomes optional:

(92) ***pčaʁe*(-xe)-m***   *a-š'ə-ʔa-ʁ*
    number-PL-OBL   3PL.IO-LOC-be-PST
    numbers   she.was.in.them
    'She has been to a number (of places).'

(93) *adəgjejə-m*   *wə-qə-ze.re-ḳʷa-ʁe-r*   ***be*(-xe)-m***   *a-ŝe*
    Adyghea-OBL   2SG.ABS-DIR-REL.FACT-go-PST-ABS   many-PL-OBL   3PL.ERG-know
    Adyghea   that.you.came   many   they.know.it
    'Many (people) know that you arrived in Adyghea.'

(94) *wəcʷəž'əqʷe bze-m*   *jə-λenəqʷe **pstewə**[-**xe**]-**r-jə** ə-wəšetə-ʁe-x*
    Uchezhuk   language-OBL   POSS-side   all-PL-ABS-ADD   3SG.ERG-test-PST-PL
    Uchezhuk   language   its.side   all   he.tested.them
    'Uchezhuk tested everyone on their knowledge of the language.'

The only quantified NPs which have plural denotation but normally do not allow plural marking with -*xe* are those containing numerals. This is illustrated in (95) for an absolutive NP with a simple numeral, in (96) for an oblique NP with a simple numeral and in (97) for an oblique NP with a complex numeral. Interestingly, however, as shown by (96)–(97), nominals containing numerals can take the oblique plural suffix -*me* (cf. Kumakhov 1971: 27–28).

(95) *ž'awš't-ew*   **mez-jə-bʁʷ** /   **mez-jə-bʁʷə-r** /   ***mez-jə-bʁʷə-xe-r**
    that.way-ADV month-LNK-nine   month-LNK-nine-ABS month-LNK-nine-PL-ABS
    thus   nine.months   nine.months   nine.months
    ḳʷa-ʁe
    go-PST
    (they).went
    'Thus passed nine months.' [E]

(96) *a-š'*   **adəge-txekʷ-jə-bʁʷə-m** /   **-txekʷ-jə-bʁʷə-me** /
    that-OBL Circassian-writer-LNK-nine-OBL -writer-LNK-nine-OBL.PL
    it   nine.Circassian.writers   nine.writers
    ***-txekʷ-jə-bʁʷə-xe-m** / ***-txekʷ-jə-bʁʷə-xe-me**   ja-txəʁe-xe-r
    -writer-LNK-nine-PL-OBL   -writer-LNK-nine-PL-OBL.PL 3PL.IO+POSS-work-PL-ABS
    nine.writers   nine.writers   their.works
    de-tə-x
    LOC-stand-PL
    they.stand.there
    'The works of nine Circassian writers are located there.'

(97) **adəge**   **ṭʷeç'ə-re**   **tfə-re-m** /   **tfə-re-me** /
    Circassian twenty-COORD five-COORD-OBL five-COORD-OBL.PL
    Circassian twenty   five   five
    ***tfə-re-xe-m** /   ***tfə-re-xe-me**   swəlṭanaʁʷe-r məsərə-m
    five-COORD-PL-OBL five-COORD-PL-OBL.PL sultanate-ABS Egypt-OBL
    five   five   sultanate   Egypt
    zə-š'-a-ʔəʁə-ʁe-r   jəλes-jə-ŝe-m
    REL.TMP-LOC-3PL.ERG-hold-PST-ABS year-LNK-hundred-OBL
    when.they.held.it.there   hundred.years
    q-je-χʷə-ʁe-ŝ …
    DIR-DAT-happen-PST-CS
    (they).rised
    'Inasmuch as more than one hundred years have passed since twenty-five Circassians ruled the sultanate in Egypt …'

A peripheral example where the suffix *-xe* appears on an NP containing a numeral is found in (98). In this example, what is pluralized is the title of a Circassian folk genre, "One hundred truths", which itself contains a numeral. Here the plurality clearly does not concern what is quantified by the numeral.

(98) *adəge-ʁaŝe-m   je-pxə-ʁe   ŝenəʁe-pstew-r-jə*
       Circassian-life-OBL DAT-connect-PST knowledge-all-ABS-ADD
       Circassian.life   related.to.it   all.knowledge
       ŝəpq-jə-ŝe-xe-m   *a-xe-b-ʁʷete-š't*
       truth-LNK-hundred-PL-OBL 3PL.IO-LOC-2SG.ERG-find-FUT
       one.hundred.truths(pl)   you.will.find.it.there
       'You will find all knowledge related to Circassian life in (the different versions of) "One hundred truths".'

Plural marking (like case marking) is infelicitous with distributive NPs containing quantifiers 'every, each', despite possible plural indexing of such NPs elsewhere (see Sections 4.1–4.2):

(99) ***ƛepq-pepč***(**-**xe-r***)   *ja-bze*   *j-e-gʷəʁə-x*
       nation-every-PL-ABS   3PL.IO+POSS-language DAT-DYN-care-PL
       every.nation   their.language   they.care.for.it
       'Every people cares for its own language.' [E]

(100) ***sabjəj-pepč***(**-**me***) *zə*   *gʷəš'ə?e* *a-txə-ʁ*
       child-every-OBL.PL   one   word   3PL.ERG-write-PST
       every.child   one   word   they.wrote.it
       'Every child wrote one word.' [E]

Definite uses. Definite NPs, whose referents are readily identifiable, are almost always overtly marked for the number of the referent: in our data, the plural form has plural denotation and the unmarked form refers to a singular individual. (101) illustrates this for an anaphoric definite common NP and (102) for a definite common NP in a bridging context:

(101) *weredə?ʷe-xe-m* *a-fe-s-ŝə-ʁa-ʁ.*   ***č'ale****(*-**xe***)**-r** *аχš'e-m*
       singer-PL-OBL   3PL.IO-BEN-1SG.ERG-do-PST-PST boy-PL-ABS   money-OBL
       singers   I.called.them   boys   money
       fa-je-x-ep
       BEN-want-PL-NEG
       (they.)do.not.want
       'I called the singers. The guys don't want money.' [E]

(102) *qʷaŝʷe-r xə-ʔʷəŝʷe-m   qə-ʔʷə-χa-ʁ.*   ****matrosə-r* / *matros-xe-r***
      boat-ABS sea-shore-OBL DIR-LOC-go.in-PST sailor-ABS   sailor-PL-ABS
      ship   shore   it.came.up.to.it   sailor   sailors
      q-je-xə-ʁe-x
      DIR-DAT-go.down-PST-PL
      they.came.down.to.it
      'The ship arrived. The sailors came down onto shore.' [E]

The picture is different for possessives and certain pronouns. Sometimes possessive phrases optionally lack the number and case markers, yet trigger plural indexing on the predicate (103)–(104). Probably, this is more common when a clause is used to convey information about the possessor (so the possessum is not individuated), but this issue needs further investigation.

(103) *çəf-gʷere   qe-s-λeʁʷə-ʁ.*   ***ə-ʔeχʷambe*(-*xe-r*)** *λə-zeč'a-ʁe-x.*
      person-some DIR-1SG.ERG-see-PST 3SG.PR-finger-PL-ABS blood-all-PST-PL
      a.person   I.saw.him/her   his/her.fingers   they.were.in.blood
      ***ə-ne*(-*xe-m*)**   *neps a-če-tə-ʁ*
      3SG.PR-eye-PL-OBL tear 3PL.IO-LOC-stand-PST
      his/her.eyes   tear it.was.under.them
      'I saw a man. His fingers were in blood. Under his eyes were tears.' [E]

(104) *t-jə-č'əle   zə λə-gʷere   de-s.*
      1PL.IO-POSS-village one man-some LOC-sit
      our.village   one a.man   s/he.lives
      ***jə-čag*(*ə-me*)** *rjen-ew   məʔerəse-jən-xe-r   q-a-pe-č'e-x.*
      POSS-tree-OBL.PL always-ADV apple-big-PL-ABS   DIR-3PL.IO-LOC-grow-PL
      his/her.trees   always   big.apples   they.grow.on.them
      'In our village there lives a man. On his trees there are always big apples.' [E]

Similar phenomena are observed with 3rd person pronouns (demonstratives, including the anaphoric pronoun *a-*) and the intensifier *jež'*.[10] These pronouns can take plural morphology (105) but occasionally appear in an unmarked form even where they trigger plural indexes and definitely refer to sets of individuals (cf. the elicited example (106) and the corpus examples (107)–(108)).

---

10 For *a-*, this has been noted by Sitimova 2004: 77–78.

(105) ***a-xe-me*** *za-ʁʷere* *w-a-xe-mə-he-me,* ***ježʼ-xe-m***
that-PL-OBL.PL once-some 2SG.ABS-3PL.IO-LOC-NEG-go.in-COND self-PL-OBL
them sometimes if.you.do.not.enter.them themselves
zečʼe *q-a-fe-ne-žʼə-ʁ-ew* *q-a-ŝʷ-e-ŝ̂ə.*
all DIR-3PL.IO-BEN-reach-RE-PST-ADV DIR-3PL.IO-MAL-DYN-do
all it.is.left.for.them it.seems.to.them
'If you do not meet with them occasionally, they (lit. themselves) think that everything is left for them.'

(106) *qa-ḳʷe-re* *maze-m* ***mə-j*** / ***a-j*** / **[mwe txəλə-m]* ***ja-ǯ!***
DIR-go-DYN month-OBL this-OBL that-OBL that book-OBL 3PL.IO+DAT-read
next month this that that book read.them
'Next month read these (books)!' [E]

(107) ***a-šʼ*** ***a****-fe-de-xe-r* *klass-pepč* *mə-z-ew,*
that-OBL 3PL.IO-BEN-correspond-PL-ABS class-every NEG-one-ADV
that similar.to.them every.class not.by.ones
mə-ṭʷ-ew *a-rə-sə-ʁe-x*
NEG-two-ADV 3PL.IO-LOC-sit-PST-PL
not.by.twos they.sat.in.them
'Such people (lit., ones like those) were in each class not just by ones and twos.'

(108) *abʒaxe-r* *z-a-šta-xe-čʼe* *bẑedəʁʷə-pšʼə-xe-m-re*
Abzakh-ABS REL.TMP-3PL.ERG-take-INC-INS Bzhedug-prince-PL-OBL-COORD
Abadzekhia when.they.took.it Bzhedug.princes
čʼemgʷəje-pšʼə-xe-m-re ***ježʼ*** ***a****-šʼə-šʼ* *pšʼə*
Temirgoi-prince-PL-OBL-COORD self 3PL.IO-LOC-be.part prince
Temirgoi.princes self being.part.of.them prince
a-f-a-ŝ̂ə-n
3PL.IO-BEN-3PL.ERG-do-MOD
they.will.make.him.for.them
'When they take Abadzekhia, the Bzhedug princes and the Temirgoi princes will choose a prince from among their own.'

Noun-headed NPs with attributive reference, i.e. with referents which have been established in discourse via a description of some sort despite the absence of a particular individual in a speaker's mind (like *the murderer of Smith* in *The murderer of Smith is presumably insane*; see Donnellan 1966) also require plural marking whenever the referent is plural (109). Still, NPs constituted by headless relative clauses in pseudoclefts, which also presumably have attributive reference, can but need not be marked as plural when referring to sets of individuals (110):

(109) ***s-jə-univ'ers'it'et-č'elejeʁaž'e*****(-*xe*)-*r*** *a* *çəf-çerəʔ*ʷ*e-xe-r* *a-rə*
1SG.IO-POSS-university-teacher-PL-ABS that person-famous-PL-ABS that-PRED
my.university.teachers that famous.people are
'My university teachers are those famous people.'[E]

(110) *qʷjeqʷe asfar jə-txəλ* *jə-nahə-be-dede*
Kuek Asfar POSS-book POSS-more-many-very
Kuek Asfar his.book his.majority
zə-f-jə-ʁe-hə-ʁe[*-xe*]*-r* *xexes-adəge-xe-r*
REL.IO-BEN-3SG.ERG-CAUS-carry-PST-PL-ABS alien-Circassian-PL-ABS
what.he.turned.them.to alien.Circassians
a-rə-x
that-PRED-PL
are
'Those to whom Asfar Kuek dedicated the majority of his books are Circassian emigrees.'

Predicative use. In West Circassian, nominals are used as predicates in constructions which are usually described as nominal predication (111) as well as in constructions employing the "adverbial" marking, including in secondary predication patterns (112) and as complements of predicates like 'consider' (113).[11]

(111) *a-d-r-jə-ṭʷə-r* *qeberteje-adəg*(*e-x*)
that-OTHER-ADJ-LNK-two-ABS Kabardian-Circassian-PL
other.two they.are.Kabardian.Circassians
'The other two are Kabardian Circassians.'

(112) *təʁʷase-re* *stud'ent-xe-r* ***vrač'***(*-x*)*-ew* *č'əle-m*
yesterday-ADJ student-PL-ABS doctor-PL-ADV village-OBL
yesterday's students as.doctors village
q-a-ʁeze-ž'ə-ʁe-x
DIR-3PL.ERG-turn-RE-PST-PL
they.returned
'Yesterday's students have returned to the village as doctors.' [E]

(113) *arab-xe-m* *adəge-xe-r* ***ǯ'wərt***(*-x*)*-ew* *a-λəte-š'tə-ʁe-x*
Arab-PL-OBL Circassian-PL-ABS Jew-PL-ADV 3PL.ERG-consider-AUX-PST-PL
Arabs Circassians as.Jews they.considered.them
'Arabs considered Circassians to be Jews.'

---

11 Because of the lack of a clear noun-verb distinction in the predicate position, the latter two types of constructions may be also described as nominal predications (cf. Vydrin 2008 for some discussion of the alleged secondary predication in West Circassian).

As these examples show, at least if the plurality of the participant described by the predicate forms is indicated elsewhere in the sentence, the appearance of plural in most of these uses is optional.

The same is true of the "intensifier" *jež'* (also regularly used as an anaphoric pronoun) – in the presence of a corresponding plural NP:

(114) *s-jə-ʔahəl-xe-m*   **jež'**(**-xe**)**-m-jə**   *a-ŝe-š'tə-r*
   1SG.IO-POSS-relative-PL-OBL self-PL-OBL-ADD 3PL.ERG-do-FUT-ABS
   my.relatives   themselves   what.they.will.do
   a-ŝe-r-ep
   3PL.ERG-know-DYN-NEG
   they.do.not.know.it
   'My relatives do not know themselves what they should do.' [E]

To conclude this section, we find that, while plural forms normally have plural reference (probably with the exception of generic and generic-like contexts), unmarked and deficient forms may be undefined for number in some contexts but not in others. Interestingly, the contexts where a form unmarked for number may have plural denotation include not only nominals with non-specific reference but also possessives and highly definite pronouns. This may be interpreted as another manifestation of the "dialectics" which makes the opposite poles of the topicality hierarchy look similar (here, a hierarchy which can be roughly formulated as DEFINITE PRONOUN >POSSESSIVE NP >DEFINITE COMMON NP > SPECIFIC NP > NON-SPECIFIC NP), as discussed by Lander (2009).

4.2 Honorific indexation

Though both Russian and Turkish, the languages with the strongest influence on Circassian, employ 2nd person plural pronouns for the expression of politeness, this has not become typical for West Circassian as spoken in Russia. Nonetheless, the language does use number contrasts for the expression of pragmatic categories.

The first-person plural possessive can be used as an honorific in certain contexts. One such context is in addressing one's mother – or any older woman – politely, as in (115). Another use of the same prefix is observed in combination with the word *λə* 'man', here with the meaning 'husband' (116). Although just one possessor is assumed, the plural is given as a sign of modesty:

(115) **t-jane**   *se* *zə-r*   *a-rə*   *sabjəj-ew* *jə-ʔe-r*
   1PL.IO-POSS+mother I one-ABS that-PRED child-ADV POSS-be-ABS
   my.mother   I   alone   am   child   whom.she.has
   'I alone am my (lit. our) mother's child.' [E]

(116) **t-jə-ƛə**   sə-ḳʷe-n-jə   z-ʁe-šxe-n
1PL.IO-POSS-man 1SG.ABS-go-MOD-ADD 1SG.ERG-CAUS-eat-MOD
my.husband   I.will.go   I.will.feed.him
'I will go and feed my (lit. our) husband.' [E]

5 Conclusions

West Circassian shows several systems of expression of number which contrast singular and plural. In the personal pronominal system, number is not an inflectional category and functions differently from non-pronominal number. For other nominals, the additive plural co-exists with associative plural and is productive. The expression of (additive) plurality, however, is optional for possessive NPs, non-locutor definite pronouns, and quantified NPs, and number need not be expressed on non-specific NPs, the formal number contrasts being weakened at the opposite poles of individuation hierarchies. Besides the systems just mentioned, West Circassian also actively exploits indexing for conveying information about number. Number marking in indexing is largely done on a semantic basis, and as a result we regularly observe that plural indexing may characterize arguments which are expressed by nominals unmarked for number.

Below we highlight some issues which we touched upon only briefly but which are definitely of interest for the typology of number and which require more research.

- Kumakhov (1971: 21–24) states that the domains of singularia and pluralia tantum are rather limited in Circassian. We do not think that these notions are applicable to languages like West Circassian at all. The morphology of West Circassian is so productive that words may be constructed in the course of speech much more easily than in many Standard Average European languages (cf. Lander & Testelets 2017: 954). If so, it is not accurate to speak of the existence and non-existence of certain forms in West Circassian. Rather, morphological productivity allows the speaker to construct any form, provided the relevant context – and indeed, in practice we do not observe any "tantum" forms in West Circassian. The system is therefore different from that of European languages.
- While languages contrasting only singular/unmarked and plural usually distinguish only between two (series of) forms, in West Circassian, we have to distinguish between three kinds of forms, namely deficient forms, plural forms (obligatorily also marked for case) and case-marked forms that are unmarked for number. This may not be a rare situation cross-linguistically if one considers deficient forms in Circassian to be a typological parallel to incorporated or pseudo-incorporated nominals (cf. Arkadiev & Testelets, 2019), i.e. nominals which are also morphologically and syntactically deficient but due to this are often

not considered a part of the number system. Taking such forms into consideration should require the revision of the typologies which only deal with binary contrasts in such cases.

–   Oblique case-marked forms may be marked for plural by two different morphemes (*-xe* and *-me*), which in most cases are interchangeable, and sometimes co-occur. Still, there are contexts where only *-me* is allowed (e.g., with numerals and with the associative plural suffix; Sections 2.3.2 and 2.3.5) or where it at least seems to be highly preferred (e.g., with some collective nouns; Section 4.1). This may suggest that the two morphemes are to be described as pertaining to different categories.

–   The fact that plural indexing is compatible with a phrase containing a distributive quantifier (Section 4) poses a problem of interpretation of the number feature. Should the semantic number of nominals and the semantic number of the corresponding indexes be considered separately, as different categories (for example, one characterizing an argument and another characterizing the situation)? Or alternatively, should we think of distributive quantifiers in West Circassian as something that does not fit into the canon of distributive quantifiers (as suggested in Arkadiev & Lander 2013)?

Acknowledgements

Abbreviations

ABS	absolutive
ADD	additive
ADJ	adjectivizer
ADV	adverbial
APL	associative plural
AUG	augmentative
AUX	auxiliary
BEN	benefactive
CAUS	causative
CNV	converb
COM	comitative
COND	conditional

COORD	coordination
CS	consecutive
DAT	dative (preverb)
DIR	directive
DISTR	distributive
DYN	dynamic
EMP	emphatic
ERG	ergative
FACT	factive relativization
FUT	future
INC	inceptive
INS	instrumental
IO	indirect object
LNK	linker
LOC	locative
MAL	malefactive
MNR	manner relativization
MOD	modal
NEG	negation
OBL	oblique
PL	plural
POSS	possessive
PR	possessor
PRED	predicative form
PST	past
PTCL	particle
Q	question
RE	refactive/reversive
REC	reciprocal
REL	relative
RFL	reflexive
RSN	reason
SG	singular
SIM	similative
TMP	temporal relativization
TRANS	translative
WIMP	weak imperative

References

Abregov, Acherdan N. 2000. *Issledovanija po leksike i slovoobrazovaniju adygejskogo jazyka* [Studies in lexicon and grammar of West Circassian]. Maykop.

Arkadiev, Peter M., Yury A. Lander, Alexander B. Letuchiy, Nina R. Sumbatova & Yakov G. Testelets. 2009. Vvedenie. Osnovnye svedenija ob adygejskom jazyke [Introduction. Basic facts on West Circassian]. In Yakov G. Testelets (ed.), *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka* [Aspects of polysynthesis: Essays in grammar of West Circassian], 17–120. Moscow: RGGU.

Arkadiev, Peter M. 2014. O nekotoryx osobennostjax sklonenija v adygskix jazykax [On some specific features of declension in Circassian languages]. In Michael A. Daniel, Ekaterina A.

Lyutikova, Vladimir A. Plungian, Sergei G. Tatevosov & Olga V. Fedorova (eds.), *Jazyk. Konstanty. Peremennye. Pamjati Aleksandra Evgenjeviča Kibrika* [Language. Constants. Variables. In memory of Alexander Kibrik], 552–563. St. Petersburg: Aleteia.

Arkadiev, Peter & Yury Lander. 2013. Non-quantificational distributive quantifiers in Besleney Kabardian. *Snippets* 27. 5–7.

Arkadiev, Peter & Yakov Testelets. 2019. Differential nominal marking in Circassian. *Studies in Language* 43(3). 715–751.

Arkhangelskiy, Timofey, Irina Bagirokova, & Yury Lander. 2018. *West Circassian (Adyghe) Corpus*. http://adyghe.web-corpora.net/ (accessed 4 December 2020).

Bagirokova, Irina G., Alina I. Baranova, & Yury A. Lander. 2020. Materialy k tipologii associativnoj množestvennosti: kubanskij dialect kabardino-čerkesskogo jazyka [Data for the typology of associative plurals: Kuban Kabardian]. *Vestnik RGGU. Ser.: Literaturovedenie. Jazykoznanie. Kul'turologija* 2020, 7, 84–106.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Daniel, Michael & Edith Moravcsik. 2005. The associative plural. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 150–153. Oxford: Oxford University Press.

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75. 281–304.

Ershova, Ksenia. 2019. *Syntactic ergativity in West Circassian*. Chicago: The University of Chicago dissertation.

Evans, Nicholas. 2006. Dyadic constructions. In Keith Brown (ed.), *Encyclopedia of language and linguistics,* 2nd edn., 24–28. Oxford: Elsevier.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Ilyina, Ekaterina D. 2011. K voprosu o reguljarnosti adygejskix form množestvennogo čisla [On regularity of West Circassian plural forms]. In *Polevye issledovanija studentov RGGU: Ètnologija, fol'kloristika, lingvistika, religiovedenie* [Field studies of students of RSUH: Ethnology, folklore studies, linguistics, religious studies], vyp. VI, 208–215. Moscow: RGGU.

Kerasheva, Zajnab I. 1982. K voprosu o kategorii čisla i kategorii opredelennosti v adygejskom jazyke [On the number category and the definiteness category in West Circassian]. *Annual of Ibero-Caucasian linguistics* IX: 111–118.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia, & Godehard Link. 1995. Genericity: An introduction. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 1–124. Chicago & London: The University of Chicago Press.

Kumakhov, Mukhadin A. 1969. Čislo i grammatika [Number and grammar]. *Voprosy jazykoznanija* 1969, 4. 65–74.

Kumakhov, Mukhadin A. 1971. *Slovoizmenenie adygskix jazykov* [Inflection in Circassian languages]. Moscow: Nauka.

Kumakhov, Mukhadin A. (ed.) 2006. *Kabardino-čerkesskij jazyk* [The Kabardian language]. Vol. 1. Nalchik: El-Fa.

Lander, Yury. 2010. Dialectics of adnominal modifiers: On concord and incorporation in nominal phrases. In Franck Floricic (ed.), *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*, 287–311. Lyon: ENS éditions.

Lander, Yury A. 2012. *Reljativizacija v polisintetičeskom jazyke: adygejskie otnositeľnye konstrukcii v tipologičeskoj perspektive* [Relativization in a polysynthetic language: West Circassian relative clauses in typological perspective]. Moscow: Institute of Oriental Studies dissertation.

Lander, Yury A. 2016. Adyghe. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation. An international handbook of the languages of Europe*. Vol. 5, 3508–3526. Berlin & Boston: de Gruyter Mouton.

Lander, Yury A. 2017. Nominal complex in West Circassian: between morphology and syntax. *Studies in Language* 41(1). 76–98.

Lander, Yury A. & Michael A. Daniel. 2019. West Circassian relative prefixes as resumptives. *Linguistics* 57(6). 1239–1270.

Lander, Yury A. & Yakov G. Testelets. 2017. Adyghe (Northwest Caucasian). In Michael Fortescue, Nicholas Evans & Marianne Mithun (eds.), *The Oxford handbook of polysynthesis*, 948–970. Oxford: Oxford University Press.

Lander, Yury, Oleg Belyaev & Irina Bagirokova. 2021. (Almost) everything is oblique in West Circassian. In Miriam Butt, Jamie Y. Findlay, & Ida Toivonen (eds.), *Proceedings of the LFG'21 Conference, On-Line*, 223–242. Stanford: CSLI Publications.

Letuchiy, Alexander. 2007. Reciprocals, reflexives, comitatives, and sociatives in Adyghe. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, 773–812. Amsterdam & Philadelphia: John Benjamins.

Nikolaeva, Liudmila. 2012. Quantifiers in Adyghe. In Edward L. Keenan & Denis A. Paperno (eds.), *Handbook of quantifiers in natural language*, 21–82. Dordrecht: Springer.

Paris, Catherine. 1989. Esquisse grammaticale du dialecte abzakh (tcherkesse occidental). In B. George Hewitt (ed.), *The indigenous languages of the Caucasus*. Vol. 2. *The North West Caucasian Languages*, 154–260. Delmar & New York: Caravan.

Rogava, Giorgi V. & Zejnab I. Kerasheva. 1966. *Grammatika adygejskogo jazyka* [A grammar of West Circassian]. Krasnodar, Majkop: Krasnodarskoe knižnoe izdatel'stvo.

Sitimova, Sara S. 2004. *Osobennosti bžedugskogo dialekta adygejskogo jazyka* [The features of the Bzhedugh dialect of West Circassian]. Maykop: The Adygheya State University dissertation.

Smeets, Rieks. 1984. *Studies in West Circassian phonology and morphology*. Leiden: The Hakuchi Press.

Smeets, Rieks. 1992. Towards an analysis of pronominal *rə* in Circassian. In Catherine Paris (ed.), *Caucasologie et mythologie compare. Actes du Colloque international du C.N.R.S. – IVe Colloque de Caucasologie (Sèvres, 27–29 juin 1988)*, 301–320. Paris: Peeters.

Sumbatova, Nina R. 2009. Kommunikativnaja struktura adygejskogo predloženija: perspektiva i focus [Information structure in West Circassian sentence: Perspective and focus]. In Yakov G. Testelets (ed.), *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka* [Aspects of polysynthesis: Essays in grammar of West Circassian], 559–611. Moscow: RGGU.

Tatevosov, Sergei. 2002. *Semantika sostavljajuščix imennoj gruppy: kvantornye slova* [Semantics of NP components: quantifiers]. Moscow: IMLI RAN.

Testelets, Yakov G. (ed.) 2009. *Aspekty polisintetizma: Očerki po grammatike adygejskogo jazyka* [Aspects of polysynthesis: Essays in grammar of West Circassian]. Moscow: RGGU.

Tutarisheva, Marzyat K. & Mariat K. Tutarisheva. 2015. K voposu o suščnosti obščego čisla v jazyke (na materiale raznostrukturnyx jazykov) [On the nature of general number in language (based on data from languages of various structures)]. *Vestnik Majkopskogo gosudarstvennogo texnologičeskogo universiteta*, 2015, 2. 120–127.

Vydrin, Arseniy. 2008. Are there depictives in Adyghe? In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary predicates in Eastern European languages and beyond*, 424–445. Oldenburg: BIS.

Yaroslavtseva, Maria V. 2017. *Voprosy vyraženija množestvennogo čisla v adygejskom jazyke* [Aspects of the expression of plural in West Crcassian]. Moscow: HSE University, School of Linguistics term paper.

III Northern Eurasia

Edward Vajda

9 Number in Ket (Yeniseian)

Abstract: Ket is a critically endangered language spoken in Siberia in the Yenisei river basin and is the last surviving member of the once widespread Yeniseian family whose typological profile is very different from its closest neighbours. Nouns and pronouns distinguish singular and plural number, usually by adding a plural suffix, with singular number left morphologically unmarked. Plural suffixes have distinct allomorphs for inanimate class nouns, kinship terms, and other animate class nouns, but there are many exceptions and irregular forms. Attributive adjectives and other modifiers are normally left unmarked for number, though a few adjectives have a plural suffix. Demonstrative pronouns, however, regularly express plurality when modifying animate class plural nouns. The Ket verb expresses agreement in singular and plural number with its subject and object and also has a variety of morphological means for expressing pluractionality, resulting in various patterns of multiple exponence of number on the verb. Among other topics, this chapter focuses on the relationship between the expression of number and animacy, which variously manifests itself in the morphology of nouns, pronouns and finite verbs. It also explains how certain irregularities in Ket number marking developed and includes comparisons with the extinct Yeniseian languages.

1 Overview

Ket is spoken today by a dwindling number of elders, mostly in remote Siberian villages near the Yenisei river or its tributaries in Turukhansk District of Krasnoyarsk Krai, the Russian Federation's second largest administrative unit. Alongside its extinct sisters – Yugh, Kott, Assan,[1] Arin, and Pumpokol – Ket belongs to the Yeniseian (Yeniseic) language family. Substrate river names and 17th century Tsarist fur tax records indicate that languages or dialects related to modern Ket were once spoken across much of south and west-central Siberia, from northern Mongolia and the forests southwest of Lake Baikal westward to the Ob river watershed and northward along the Yenisei to the Arctic Circle (Map 1).

Three Ket dialects have survived into the early 21st century: Southern Ket (spoken in several villages, including Kellog, Sulomai, and Alinskoye), Central Ket (Sur-

---

[1] In terms of linguistic similarity, Assan could easily be considered a dialect of Kott. However, as in the case of Yugh, the speakers considered themselves to be ethnically distinct. For this reason, Assan and Kott, like Ket and Yugh, are generally regarded today as closely related pairs of languages.
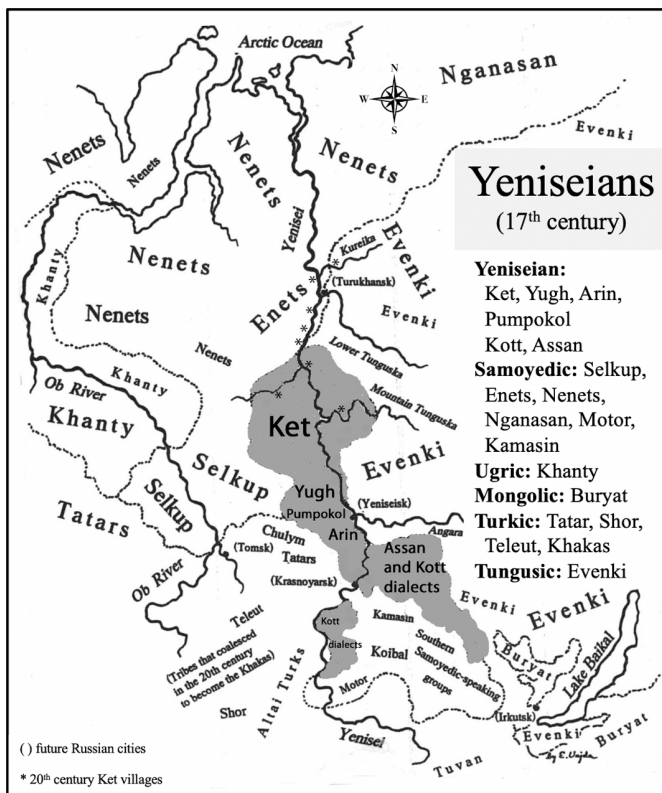
Map 1: Language distribution in central Siberia during the 17th century.

gutikha and Baklanikha), and Northern Ket (Maduika and Kureika). At the time of this writing, these dialects together have no more than a few dozen fully fluent native speakers, all over the age of 65. There are only small differences in how each Ket dialect expresses number, and most are of a predictable phonological nature. Except where otherwise noted, this chapter cites Southern Ket forms. Number marking in Ket broadly resembles Yugh (Werner 1997a), a distinct but closely related language that disappeared in the 1970s. Most 20th century Soviet scholarship on Yeniseian languages treated Yugh as another Ket dialect, calling it "Sym-Ket" after a river where most speakers lived. By contrast, Southern, Central and Northern Ket were called "Imbat Ket" after *Inbak* – a 19th century ethno-geographic term for several downriver groups of Kets. The other documented Yeniseian language varieties, all of which disappeared between 1730 and 1850, are more distantly related and sometimes provide deeper insights into the historical development of number marking in the family. The extant material representing the Kott language, which belongs to a different primary branch than Ket and Yugh, is the most useful in this regard because it contains many inflected word forms, including plurals, thanks to a gram-

mar sketch and dictionary recorded by the Finnish field linguist M. A. Castrén (1858).

Yeniseian languages are in many ways structurally unlike the "Ural-Altaic" belt of language families originally spoken across nearly all of the remaining portions of northern and Inner Eurasia. The most striking difference is the Ket polysynthetic finite verb, which is based on a template of ten morpheme classes, most of which, at least historically speaking, are represented in prefixal positions. By contrast, verbs in Turkic, Tungusic, and Uralic languages are exclusively suffixing in their inflectional morphology. Possession in Yeniseian is likewise expressed using markers placed in front of the possessed noun (or noun phrase), rather than by suffixes as in the neighboring languages.

This chapter covers all aspects of how Southern Ket expresses number categories. It also considers the diachronic origins – whether genealogical or contact-induced – behind the various formal systems of number marking across the Yeniseian family. In keeping with the volume's general structure, section 2 provides descriptions of number marking across the language's major form classes. Subsection 2.2 examines the Ket pronoun system, which manifests all of the language's core number distinctions in one way or another. It also provides a general overview of how animacy intersects with plural marking patterns in several areas of the morphosyntax. Subsection 2.3 investigates the complex motivations behind how plural suffix allomorphs are distributed across the noun lexicon, and also looks at sporadic examples of dual number marking that arose through morphological reanalysis of certain noun stems. Subsection 2.4 describes morphological techniques used for expressing pluractionality. Section 3 discusses agreement morphology. Subsection 3.1 examines number agreement in demonstrative pronouns and adjectives. Subsection 3.2 examines how adjective and action nominal suffixes have been reanalyzed as plural markers in a number of stems. Subsection 3.3 discusses the number agreement suffixes found on certain lexical classes of subject complements in the predicates of clauses with no finite verb form. Finally, subsection 3.4 turns to the highly complex polysynthetic Ket verb to explain its intricate system of class and number agreement with subjects and objects. Section 4 remarks on what is known about the discourse functions of number marking. Section 5 summarizes the typology of number marking in modern Ket and attempts to sort out which techniques were inherited from Proto-Yeniseian and which were later innovated. The discussion throughout this chapter also points out noteworthy features and categories related to number that are found in other languages but absent in Yeniseian.

Much of the material here was covered earlier in Porotova (1990), the only existing monograph devoted to the expression of plurality in Yeniseian. Accessible treatments of most aspects of Ket number marking can also be found in Werner (1997b), Vajda (2004), Georg (2007), and Nefedov & Vajda (2015).

2 Pronominal, nominal, and verbal number

2.1 Generalities

Number marking features shared widely with other Siberian language families include the productive use of suffixes to mark plurality in nouns. The expression of number in Ket and other Yeniseian languages is based on a formal distinction between singular (one) and plural (two or more), and this dichotomy strongly affects the morphosyntax of nouns, pronouns, and finite verbs. Yeniseian also has a noun class (grammatical gender) system based on the contrast between inanimate and animate class, with animate class further subdivided into masculine and feminine gender. Ket noun class (animacy and gender) interacts with number marking in several important ways, with inanimate class nouns showing an overt distinction between singular and plural only in the noun form itself, but not in the possessive markers or verb agreement affixes that cross-reference them. Finally, modern Ket has also developed a typologically unusual system of phonemic tones based on an amalgam of features involving melody, length and phonation type.[2] Tone is marginally germane to the topic of number in Ket because a small irregular set of singular vs. plural noun forms are formally distinguished solely by tonal differences.

2.2 Pronominal number

2.2.1 Independent pronouns

The system of personal and anaphoric pronouns in Ket offers a good starting point for discussing number marking. These pronouns convey all of the core number distinctions present in Yeniseian morphosyntax (singular vs. plural) and provide clear examples of the most typical formal means used to mark them (plurality expressed

---

2 The Southern Ket data in this chapter derive from the author's own fieldwork. Monosyllabic phonological words distinguish four phonemic tones. These are transcribed using a macron for high-even tone (*sūl* 'blood'), apostrophe for rising laryngealized tone (*su'l* 'white salmon'), double vowel letter for rising-falling tone on a geminate vowel (*suul* 'snow sled'), and grave accent for falling tone (*sùl* 'cradle hook'). Most polysyllabic phonological words have instead an accent-like pitch on the first syllable, as in *qópqun* 'cuckoo bird'. In a much smaller number of polysyllabic words the pitch peak falls on the second syllable, one example being the plural form *qopqún* 'cuckoo birds'. Because syllable-initial pitch in polysyllables is far more common, it is left unmarked except in disyllabic noun stems like *qópqun* 'cuckoo bird', where it is included to call attention to its relevance for number marking. Finally, each of the three Ket mid vowel phonemes have two regular allophones. The symbol /ə/ is used to transcribe the mid back unrounded vowel, which is realized phonetically as [ɣ̌·] under high-even tone and as [ʌ] elsewhere. In a similar fashion, /o/ transcribes both [ō·] and [ɔ] and /e/ transcribes [ē·] and [ɛ].

through suffixes that normally contain a nasal consonant, either /n/ or /ŋ/). Table 1 shows the system of singular and plural animate personal pronouns in Southern Ket:

Tab. 1: Ket personal and anaphoric pronouns.

1SG *ād*	1PL *ətn*
2SG *ū*	2PL *əkŋ*
3SG.ANIM *bū*	3PL.ANIM *būŋ*
	3INAN *tude*

The 3rd person anaphoric pronouns *bū* 'he / she' and *būŋ* '(animate class) they' normally refer back to nouns denoting human beings, but sometimes animals as well. Other nouns are generally replaced anaphorically by *tude* 'that (one)', 'those ones' (pronounced [tuɾɛ] in Southern Ket), which can also be used adnominally as a demonstrative pronoun.[3]

The simplicity of this paradigm stems in part from the fact that Yeniseian pronouns do not distinguish dual number. Nor is there any inclusive/exclusive distinction. The 1st person plural pronoun *ətn* 'we' can mean 'you and me' or 'you and us' (inclusive); or it can mean 'me with another or others' (exclusive). In contrast to many modern Indo-European languages, the Ket 2nd person plural pronoun *əkŋ* 'you' was not traditionally used to address a single individual to show respect or mark social distance, though recent Russian interference influenced such occasional usage among the last generation of speakers.

Ket pronoun forms also exemplify the number marking strategy that is overwhelmingly favored across Yeniseian nominal morphology. The core distinction between singular (one) and plural (more than one) is formally expressed by augmenting the singular form with a suffix containing either the velar nasal /ŋ/ or alveolar nasal /n/. There is no grammatical means for marking singular number, aside from the absence of plural marking. Plural suffixation in nouns sometimes entails irregular morphophonemic changes, as seems to be the case with Ket 1st and 2nd plural pronouns as well. Finally, although Yeniseian morphosyntax lacks a grammatical category of dual number, there are sporadic instances where a nasal coda in certain noun stems denoting paired objects such as mittens has been reanalyzed as a dual marker (see Section 4.3).

The personal pronoun system likewise affords a convenient introduction to the important role played by animacy in Ket morphosyntax. We have already seen that

---

3 Possessive markers are considered clitics because they normally attach to the preceding word if one is available rather than to the following possessum noun, as shown here. Substitutions of 3rd person animate class *bū* for inanimate class *tude* with reference to inanimate class entities have occasionally been documented, possibly due to Russian interference.

considerations of animacy determine which third-person anaphoric pronouns are used. Third-person singular *bū* 'he / she / him / her' and plural *būŋ* 'they / them' normally reference animate entities, while the demonstrative form *tude* 'that' is used when referring back to singular or plural inanimate class entities. The biological gender distinction in third-person animate singular *bū* is covert in the pronoun itself, but surfaces elsewhere in the clause by triggering different masculine and feminine forms of singular possessive clitics (1), subject concord suffixes on predicate adjectives (2), and subject/object agreement affixes in the finite verb complex (3):

(1) Animate singular gender distinction in possessive marking: (a) masculine, (b) feminine

  a. *bu=da⁴*   *qu's*
    3SG.PRON=M.POSS tent
    'his tent'

  b. *bu=d*   *qu's*
    3SG.PRON=F.POSS tent
    'her tent'

(2) Animate singular gender distinction in predicate adjectives: (a) masculine, (b) feminine

  a. *hīk* [hīˑɣ] *sel-du*
    man   bad-M.SBJ
    'The man is bad.'

  b. *qīm*   *sel-da*
    woman bad-F.SBJ
    'The woman is bad.'

(3) Animate singular gender distinction in verb-internal subject/object agreement:
  a. masculine subject with feminine object,
    hīk *qīm*   *d-i-toŋ*
    man woman 3M.SBJ-3.OBJ-see
    'The man sees the woman.'

---

4 The combination *bu=da* is pronounced [buɾa] in Southern Ket. To keep the underlying forms more transparent, the transcription ignores the allophonic intervocalic lenition of /d/ to [ɾ], /b/ to [v], /k/ to [ɣ] and /q/ to [ʁ], as well as voicing of /k/ to [g] next to a voiced consonant. It also ignores word-final devoicing of /d/ to [t] and /b/ to [p]. Note that word-final /p/ in words such as *dāp* 'shoulder' belongs to the same phoneme as word-initial /h/ and not to /b/. The symbols /p/ and /h/ here could have been unified by choosing one or the other in all cases; however, doing so would have given a misleading impression of how these allophones are actually pronounced in modern Ket. Phonetic forms are sometimes provided in square brackets, especially in less transparent cases such as certain Southern Ket high-even tone words and all falling-tone words, where

b.   feminine subject with masculine object
    qīm   hīk   da=a-toŋ
    woman man 3 F.SBJ=3M.OBJ-see
    'The woman sees the man.'[5]

The agreement contrast triggered by masculine and feminine animate singular forms does not extend to expressions of plurality. However, the difference between inanimate as opposed to animate entities is regularly maintained in the same three environments of possessive constructions (4), subject concord (5), and verb-internal subject or object agreement (6):

(4)   3P inanimate (a) and animate (b) possessive clitics
    a. *tu-de=d   kīd*
       that-INAN=INAN.POSS price
       'its price' / 'their.INAN price'

    b. *buŋ=na   kīd*
       they=ANIM.PL.POSS price
       'their.ANIM price'

(5)   3P plural inanimate (a) and animate plural (b) subject concord in predicate adjectives
    a. *tu-de   sel-am*
       that-INAN bad-3INAN.SBJ
       'It is bad.' / They.INAN are bad.'

    b. *būŋ   sel-aŋ*
       they.ANIM bad-3ANIM.PL.SBJ
       'They.ANIM are bad.'

(6)   Verb-internal agreement with inanimate (a) and animate (b) plural objects
    a. *būŋ   tu-de   du-Ø-toŋ-n*
       they.ANIM that-INAN 3ANIM.SBJ-3INAN.OBJ-see-ANIM.PL.SBJ
       'They see it.' / 'They see them.INAN.'

---

intervocalic lenition persists even though the final vowel has elided: *hīk* [hī·ɣ] 'man', *qòq* [ᵠχɔ̀ʁ] 'star'.

5 The feminine class subject marker /da/ is also a clitic that prefers to attach to a preceding word, whenever one is available, just like the possessive markers described earlier. Enclisis is less likely when the preceding word ends in an obstruent, as in this example. All Ket morphemes marked as clitics in this chapter behave in a similar way.

b. *bӯŋ   de'ŋ   d-aŋ-toŋ-n*
   they.ANIM people 3SBJ-3ANIM.PL.OBJ-see-ANIM.PL.SBJ
   'They see the people.'

Examples (1) to (6) provide a good introduction to how animacy and gender (noun class) interacts with the expression of number across Ket morphosyntax. Only singular referents trigger the overt expression of gender, while the contrast between animate and inanimate extends to patterns of plural agreement, as well. The pervasive role of animacy in Ket number marking is further explored in the next section.

2.2.2 Number in possessive constructions

One key Yeniseian typological feature absent from other families of northern Eurasia outside of Indo-European is the presence of a noun class system, examples of which have already been provided above. Unlike Russian and other Indo-European languages with grammatical gender, however, Ket agreement classes reflect foremost a distinction between animate and inanimate entities, with animate class nouns secondarily divided into masculine and feminine sub classes in singular forms – a division of much less significance to the grammar overall. The animate superclass subsumes not only humans, animals, birds, fish, reptiles, amphibians, and insects, but also certain body parts, as well as trees species. A few animate class nouns denote culturally salient inanimate objects such as pinecones, tent poles, buttons, and cradle hooks, with the masculine designation typically favoring objects of special positive value in traditional Ket lifeways. Grammatical animacy and gender are usually covert in the noun stem itself.

The class division between inanimate and animate has broad ramifications for number marking in several domains of Yeniseian morphosyntax. The next several paragraphs analyze the role of animacy and gender in the shape of possessive markers and various word forms that evolved on the basis of possessed noun constructions.

The expression of number in possessive markers intersects with animacy and gender distinctions. The forms in Table 2 illustrate how this pattern is reflected in the possessive forms of Ket personal pronouns:

Tab. 2: Possessive forms of personal pronouns.

1SG *a=b*	1PL *ətn=na*
2SG *u=k*	2PL *əkŋ=na*
3SG.M *bu=da*	3PL.ANIM *bӯŋ=na*
3SG.F *bu=d*	
3INAN *tu-de=d*	

Singular possessive markers, shown as enclitics in Table 2, distinguish person and gender of the possessor, except that the 3rd person feminine animate singular and inanimate markers formally coincide. More importantly, the same inanimate class possessive clitic is used for singular and plural number. Clitics only differentiate masculine and feminine and person categories with regard to singular entities, while the plural possessive markers distinguish animacy only. In the absence of a preceding pronoun to disambiguate person, constructions with an animate plural possessor such as *na=ām* could therefore be understood to mean either 'their mother', 'our mother', or 'your.PL mother'. In cases where no possessor is named, the possessive marker can attach either to the following word as proclitic or to the preceding word as enclitic, depending on factors such as sentence structure and intonation. Enclisis is favored whenever a preceding word is available as host, particularly those that end in a vowel or sonorant:

(7) Possessive marker as (a) proclitic or (b) enclitic
　　a. *da=ām*
　　　his=mother
　　　'his mother'

　　b. *əqaj=da　　am-as　　　d-ol-daq*
　　　formerly=his　mother-with　3M.SGJ-PST-live
　　　'He used to live with his mother.'

Several types of Ket morphological structures derive from possessed nouns, including some case suffixes, most postpositional constructions, and a minor lexical class known as directional adverbs. All of them retain the same interplay between noun class and the expression of number displayed in possessed noun constructions.

　　Three Southern Ket case suffixes – dative (to, toward), adessive (at, in, having), and ablative (from, out of)[6] – are preceded by a possessive augment expressing the same characteristic gender and animacy distinctions as those associated with possessed nouns. When animate class plural nouns are used with these case suffixes, the resulting word forms exhibit multi-site plural marking:

---

6 The so-called benefactive case forms of *-dita* 'feminine singular or inanimate', *-data* 'masculine singular', and *-nata* 'animate plural', appear to be idiolectal reductions of adessive case forms.

(8) Adessive (a–b), dative (c–d), and ablative (e–f) case suffixes

i. with inanimate nouns

a. *suul-aŋ-di-ŋten*
   snow.sled-PL-INAN-ADES
   'in the snow sleds'

c. *baŋ-n-di-ŋa*
   land-PL-INAN-DAT
   'to our (your, their) lands'

e. *bes-n-di-ŋal*
   rabbit-PL-INAN-ABL
   '(made) out of rabbit pelts'[7]

ii. with animate plural nouns

b. *am-aŋ-na-ŋten*
   mother-PL-ANIM.PL-ADES
   'at the mothers'

d. *hun-aŋ-na-ŋa*
   daughter-PL-ANIM.PL-DAT
   'to our (your, their) daughters'

f. *bes-n-na-ŋal*
   rabbit-PL-ANIM.PL-ABL
   '(motion) from rabbits (living animals)'

The markers *-di-* and *-na-* in (8) are glossed simply as INAN and ANIM.PL, respectively, since the earlier possessive meaning was lost. Vajda (2013b: 80–84) has argued that the segment /ŋ/ located at the beginning of the modern dative, adessive and ablative case suffixes is a vestige of an ancient possessive connector once present in all possessive constructions. Regardless of whether this interpretation is valid, structures such as (8a), (8b), and (8c) suggest that the three possessive-augmented case suffixes may derive from underlying noun forms or postpositions, though the precise etymologies remain unclear. The other case suffixes – locative, comitative-instrumental (with), prolative (along or though), and caritive (without) – attach to nouns or pronouns without a preceding possessive augment, for reasons that are likewise unclear.

Ket postpositions, most of which require a case suffix, form complex constructions that are often preceded by a possessive marker that agrees in number, animacy and gender with the preceding noun or pronoun stem; they preserve all of the same morphosyntactic distinctions expressed in true possessive constructions. Some Ket postpositions transparently derive from body-part nouns, suggesting that postpositional constructions arose via the grammaticalization of possessed nouns. For example, the Southern Ket postposition *-un-* 'below', is cognate with a noun root meaning 'belly' or 'underside' in the extinct Yeniseian languages. In Ket it no longer expresses its original anatomical meaning, having been replaced by *hūj* 'belly', though Ket does preserve a cognate noun *u'n* 'sled runner'. The examples below use inanimate and animate class clitics to show the homology of the structures displayed by possessed nouns (9) and by the postpositional constructions grammaticalized on the basis of possessed nouns (10):

---

7 See Werner (1997b: 107) for the full paradigms of animate class *be's* 'rabbit (live animal)' vs. inanimate class *be's* 'rabbit (pelt)'. The inanimate class compound *besiŋolt* 'rabbit pelt' (with *iŋolt* 'pelt', 'fur') also exists.

(9) Possessed nouns with case suffixes

    a. *suul=d*   *u'n*

       snow.sled=INAN.POSS runner (< underside)

       'sled runner (one of two long pieces of wood beneath a sled for gliding on snow)'

    b. *ked=da*   *hūj*

       person=M.POSS belly

       'person's belly'

(10) Contrast between (a) possessed noun construction and (b) postpositional construction

    a. *ked=da=huj-di-ŋal*

       person=M.POSS=belly-INAN-ABL

       'from the person's belly'

    b. *ked=da=huj=d=un-di-ŋal*

       person=M.POSS=belly-INAN.POSS=under (< underside)-INAN-ABL

       'out from under the person's belly

Directional adverbs, which express orientation relative to a specific point in space, are built exactly like possessed noun constructions. The two most common directional roots are *ikda* [igda] 'downland', 'orientation from forest to river' and *aka* [aɣa] 'upland', 'orientation from river to forest'. Directional word forms are adverbs that consist of a possessive clitic and directional root, followed by a case suffix. If the required case is ablative or dative, the case suffix must be preceded by the connector *-di-*, originally the inanimate class possessive marker (the adessive case is not used in directional constructions).

(11) Directional adverbs

    a. *assanos-in na=ikda-di-ŋal*

       hunter-PL ANIM.PL.POSS=downland-INAN.POSS-ABL

       '(movement) from a location downland with respect to the hunters'

    b. *qu-ŋ*   *d=aka-di-ŋa*

       tent-PL INAN.POSS=upland-INAN.POSS-DAT

       '(movement) toward a location upland with respect to the tents'

Ket possessed nouns, certain case forms, postpositional constructions, and directional adverbs all display multi-site marking of number, animacy, and what is often called "gender" (masculine vs. feminine animate singular agreement). Instances of multiple exponence in all of these structures arose through the coalescence of formerly discrete words into single morphological forms. The expression of noun class in connection with number in certain case forms and postpositional constructions

is unique among the indigenous language families of landlocked northern Asia. The feature of multi-site plural marking in some of these word forms (cf. 8 above) is likewise unique to Yeniseian across this broader geographic area.

The effect of animacy on number marking patterns extends to other parts of Ket morphosyntax, as well. Section 2.3 explores how considerations of both referential and grammatical animacy influences the choice of plural suffix allomorphs in Ket noun forms. Section 3.3 analyzes the effect of animacy on the subject concord suffixes required by certain subject complements in linking verb clauses, which lack a finite verb. Section 3.4.1 examines the role played by noun class in assigning subject and object number agreement markers in finite verb forms.

2.3 Nominal number

Yeniseian nouns have no grammatical marker of singular number. Count nouns typically distinguish number by augmenting the bare singular stem with a plural suffix containing a nasal consonant – either /ŋ/ or /n/. There are no nominal plural prefixes. Stem reduplication is absent as a means of expressing plurality or pluractionality, being limited in Yeniseian morphology to a few expressive interjections. Overall, the choice of noun plural suffix shape follows several broad patterns that together form a complex interweave of semantics (animate vs. inanimate class) with certain formal features of stem structure, the most important being whether the noun stem already ends in /ŋ/ or /n/. There are also many irregularities caused by diachronic processes of reductions in the noun stem's final syllable or coda, some of which have caused the loss of the plural suffix itself. The subsections below discuss the motivations that determine plural noun forms (2.2.1), dual number marking arising sporadically through reanalysis of noun stem coda nasals (2.2.2), as well as *pluralia tantum*, *singularia tantum*, singulatives, and a few other number-related lexical categories at work in Ket noun morphology (2.2.3). This section goes into considerable detail because the expression of plurality in nouns is fundamental to number marking throughout Yeniseian morphology, given that nearly all other number marking categories reflect the singular/plural dichotomy inherent to nominal morphology.

2.3.1 The morphology of noun pluralization

Animacy plays a fundamental role in how nouns build their plurals in Yeniseian languages. Animate and inanimate class nouns typically have different plural suffix shapes, except where phonological or diachronic factors intervene. Inanimate class nouns normally take the suffix *-(V)ŋ*. The symbol *(V)* here and elsewhere represents a phonologically undetermined vowel that is often inserted between stem coda and nasal suffix:

(12) Examples of inanimate class noun pluralization using *-(V)ŋ*

qobad	'back'	*qobad-aŋ*	'backs'
bətl	'bubble'	*bətl-aŋ*	'bubbles'
suʼk	'bowl'	*suk-ŋ*	'bowls
qoŋloq	'bell'	*qoŋloq-ŋ*	'bells'
asl	'ski'	*asl-iŋ*	'skis'[8]
àj	'bag'	*aj-éŋ*	'bags'

The animate class noun plural suffix is *-(V)n*, with alveolar rather than velar nasal. It too is sometimes augmented by a vowel after a stem-final consonant:

(13) Examples of animate class noun pluralization using *-(V)n*

hīk	'man'	*hik-n*	'men'
ostúk	'Ket'	*ostúk-ən*	'Kets'
beʼs	'rabbit'	*bes-n*	'rabbits'
qīm	'woman'	*qim-n*	'women'
hə́mka	'Evenki'	*hə́mka-n*	'Evenkis'
laʼt	'beaver'	*lat-n*	'beavers'

The distribution of the alveolar as opposed to velar nasal consonant in noun plural suffixes is determined by much more than grammatical animacy, however. In fact, this deceptively elegant dichotomy has so many exceptions that recognizing a noun's inherent class membership based on its surface plural form is unreliable. (Noun class can be detected more reliably from observing the form of possessive markers or agreement morphology). The quality of the vowel before either of these basic noun plural suffixes is often unpredictable, as well. Some of these vowels appear to be epenthetic, which often appears true for [ə] or [ɯ]. Others reflect preservation of an original stem vowel lost in the singular form, as tends to be true of many plurals that augment the nasal suffix using [i], [e], or [a]. In other cases, it remains unclear what determines the choice of vowel in noun plural suffixes (or even whether any vowel appears at all).

One fairly consistent semantic exception to the basic pattern of animate plurals marked with *-(V)n* involves nouns denoting kinship. Despite obviously belonging to the animate class, most kinship terms build their plural by adding the suffix form *-aŋ*:

(14) Kinship nouns pluralized with *-aŋ*

ām	'mother'	*am-áŋ*	'mothers'
ōb	'father'	*ob-áŋ*	'fathers'
huʼb	'son'	*hub-aŋ*	'sons'
huʼn	'daughter'	*hun-aŋ*	'daughters'

---

8 The nouns *bətl* 'bubble' and *asl* 'ski' both end in a syllabic sonorant and thus are disyllables with rising tone on the initial syllable.

Full vowel forms of the kin noun plural suffix are consistently found in other Yeniseian languages, as well, though not necessarily with the same vowel (cf. Yugh *amén* 'mothers', *ob-éŋ* 'fathers'. Given that this vowel usually draws the pitch peak onto itself in disyllables, the Ket kinship plural suffix *-aŋ* is probably etymologically distinct from inanimate class *-(V)ŋ*, which does not normally attract the accent. However, as with every general pattern of Ket noun plural formation, kin nouns also include a few exceptional plural forms. Some nouns that express kinship take the regular animate class suffix *-(V)n* instead of the kin noun plural suffix *-aŋ*. Two examples are *qīm* 'wife, woman' → *qim-n* 'wives, women' and *tēd* 'husband' → *tad-n* 'husbands', which do not seem to be treated as primary kin terms, at least as regards the morphology of plural formation. Conversely, a few animate class nouns that are not kinship-related take *-aŋ* rather than the expected suffix *-(V)n*: *ēs* 'god', 'deity' → *esáŋ* 'gods', 'deities'; *qān* 'khan' → *qanáŋ* 'khans'. The fact that the suffix vowel attracts the pitch peak in the latter two plurals suggests that they share the same suffix *-aŋ* used by most kinship nouns. Most Ket personal pronouns (shown earlier in Table 1) also use the velar nasal as a plural ending. It should be mentioned here that Ket lacks dyadic kinship forms altogether.

Nouns in (15) take the regular inanimate class suffix *-(V)ŋ* yet trigger masculine animate or feminine animate agreement. This pattern occurs with some animate nouns that do not denote people or animals, and is observed with certain body part nouns, nouns denoting tools or other important objects, and some nouns denoting lower forms of life, which tend to belong to the feminine subclass:

(15) Grammatically animate class nouns that pluralize using inanimate class *-(V)ŋ*
       Body part nouns that are grammatically (m) masculine or (f) feminine animate

buʼs	'penis' (m)	*bus-aŋ*	'penises'
huud	'tail' (f)	*hud-aŋ*	'tails'
tūl	'navel' (f)	*tul-aŋ*	'navels'

Other inanimate objects that are grammatically (m) masculine or (f) feminine animate

quu	'pole' (m)	*quu-ŋ*	'poles'
álal	'(boat) seat' (f)	*álal-aŋ*	'(boat) seats'

Lower forms of life that are grammatically feminine animate (f)

biilt	'martin' (f)	*bilt-aŋ*	'martins'
hankó	'toadstool' (f)	*hankó-ŋ*	'toadstools'
ulól	'leech' (f)	*ulól-aŋ*	'leeches'

Although such nouns are grammatically animate as indicated by anaphoric pronoun usage, possessive morphology, and subject or object agreement, most take the

inanimate class plural suffix *-(V)ŋ* rather than animate class *-(V)n*, except where additional phonological factors such as nasal dissimilation play a role (see below, this subsection). In this case, plural formation is governed by referential considerations of animacy, whereas nominal categorization as reflected in agreement patterns is determined on a lexical basis.

There is also a tendency for recent loanwords to take the "inanimate class" plural suffix *-(V)ŋ*, even where such nouns denote human beings or animals:

(16) Examples of animate class loanwords pluralized using *-(V)ŋ*.

bótaj	'rich man'	(< Russian *bogatij* 'rich')	*bótaj-aŋ*	'rich men'
ópsa	'sheep'	(< Russian *ovtsa* id.)	*ópsa-ŋ*	'sheep.PL'
mína	'pig'	(< Russian *svin'ja* id.)	*mína-ŋ*	'pigs'
ondátur	'muskrat'	(< Russian *ondatr* id.)	*ondátur-aŋ*	'muskrats'

This pattern seems to have become relatively productive in the last phase of Ket language usage, though exceptions can be found, such as *koska* 'cat' → *koska-n* 'cats', where animate *-(V)n* pluralizes a 20th century borrowing of Russian *koška* 'cat'. In 2008, during a conversation with her mother, fluent speaker V. A. Romanenkova spontaneously used the Russian noun *sobolʲ* 'sable' in the hybrid Russian/Ket plural form *sobolʲ-aŋ* in place of native Ket *et-n* 'sables' with its canonical animate class suffix *-(V)n*. The common usage of *-(V)ŋ* to pluralize recent loanwords regardless of their meaning perhaps underscores a more general trend away from the originally fundamental morphosyntactic dichotomy in Yeniseian languages between animate and inanimate class entities.

The system of Ket noun plural formation also involves numerous phonological and morphological exceptions that override the underlying triple contrast between generic animate *-(V)n* vs. kinship *-aŋ* vs. inanimate *-(V)ŋ* described above. The most regular among these patterns involves dissimilation of the plural suffix nasal consonant in stems that end in either /n/ or /ŋ/, and sometimes /m/. Noun stems ending in /n/ nearly always take *-(V)ŋ* even when they belong to the animate class:

(17) Pluralization of animate class nouns ending in /n/

bústen	'wasp'	*bústen-aŋ*	'wasps'
tuln	'lizard'	*tuln-eŋ*	'lizard'
lūn	'grayling (fish)'	*lun-áŋ*	'graylings'

Exceptions such as *kùn* 'wolverine', *kun-en* ~ *kunn* 'wolverines' probably involve the diachronic loss of a final consonant other than /n/ that originally occupied the stem-final position.

Through an analogous process of dissimilation, inanimate class noun stems ending in either /ŋ/ or /m/ nearly always require the plural suffix *-(V)n* even where *-(V)ŋ* would be expected.

(18) Pluralization of inanimate class noun stems ending in /ŋ/ or /m/

ka'ŋ	'hole'	*kaŋ-en*	'holes'
hə'ŋ	'fishing net'	*həŋ-en*	'fishing nets'
àŋ	'rope'	*aŋ-en*	'ropes'
qām	'arrow'	*qam-en*	'arrows'
kulém	'lid'	*kulém-n*	'lids'

One exception is *bōŋ* 'corpse' → *boŋ-áŋ* 'corpses', where dissimilation fails to occur for unknown reasons. This noun belongs to the feminine animate class in terms of agreement, but even if the velar nasal was chosen through semantic influence from the word's lack of logical animacy, the suffix should still undergo nasal dissimilation to *-an*, but it does not. Instead, it seems to pattern formally more like a kinship noun, including the syllable-final accent shift.

The kin noun plural suffix *-aŋ* is likewise never affected by nasal dissimilation, and its velar nasal appears even after /m/: *ádbam* 'sister-in-law' → *ádbam-aŋ* 'sisters-in-law'. It is also noteworthy that other animate class nouns ending in /m/ preserve the canonical animate class suffix form *-(V)n*: *tēm* 'goose' → *tem-n* 'geese'. The stem coda /m/ has no effect on the form of the animate class plural suffix *-(V)n* either.

A different process, possibly historically a type of assimilation, replaces the inanimate class suffix *-(V)ŋ* with *-(V)n* in many though not all nouns with stems ending in the segment /s/:

(19) Inanimate class noun stems ending in /s/ pluralized with *-(V)n* rather than *-(V)ŋ*

kūs	'ringworm'	*kus-en*	'ringworms'
qāks	'wound'	*qaks-en*	'wounds'
ádes	'iron nail'	*ádes-n*	'iron nails'
qóles	'hoof'	*qóles-n*	'hooves'
də́les	'willow thicket'	*də́les-n*	'willow thickets'

Other inanimate class nouns with stems ending in /s/, however, form their plurals with the expected velar nasal, sometimes with loss of /s/, rather than by assimilating the nasal plural segment to alveolar /n/: *qu's* 'birchbark tent' → *qu'ŋ* 'birchbark tents', *tu's* 'rock' → *tə'ŋ* 'rocks'. Finally, the velar nasal is doubly unexpected in the plural form of the masculine animate class noun *ùs* 'birch tree' → *us-eŋ* 'birch trees' since the animate class suffix *-(V)n* would have been expected on the grounds of its animacy. Even if inanimate class *-(V)ŋ* appears in this plural because it does not denote a person or animal, the lack of /ŋ/ to /n/ assimilation after stem-final /s/ remains unaccounted for. The same observations are germane in the case of the plural form of masculine animate class *bu's* → *bus-aŋ* 'penises'. The reasons behind all of these phenomena remain at present unelucidated.

A far more frequent and completely regular departure from the underlying semantic patterning of Yeniseian plural allomorphs affects any noun stem built by adding the universal nominalizing suffix *-s*. All of these stems are pluralized with the suffix form *-in*. The vowel /i/ derives from the historic form of the nominalizing suffix *-si, which has eroded to *-s* in singular stems. Virtually any non-noun form, including postpositional constructions and inflected finite verbs, can be nominalized by adding the word-final suffix *-s*. The plural of such nominalizations always ends *-in*, regardless of animacy considerations.

(20) Use of *-in* to pluralize nominalizations in *-s*

    a. *aq-s*   *aq-s-in*
        rot-NMLZ   rot-NMLZ-PL
        'something rotten'   'things that are rotten'

    b. *d=b-il-bet-in-s*   *d=b-il-bet-in-s-in*
        3ANIM.SBJ=INAN.OBJ-PST-make-   3ANIM.SBJ=INAN.OBJ-PST-make-
        ANIM.PL.SBJ-NMLZ   ANIM.PL.SBJ-NMLZ-PL
        'something that they made'   'things that they made'

    c. *ul-di-ŋal-s*   *ul-di-ŋal-s-in*
        water-INAN-ABL-NMLZ   water-INAN-ABL-NMLZ-PL
        'the one (pulled) from the water'   'the ones (pulled) from the water'

There are no exceptions to this highly productive pattern. In example (23c) the plural suffix *-in* would be used regardless of whether what emerged from the water was a person, animal, or inanimate object.

    Other nouns undergo an array of irregular changes when a plural suffix is added. Their full range is too broad to be covered exhaustively here since each exceptional form probably has its own idiosyncratic historical explanation. The difference between the plural vs. singular stem phonology in some of these number pairs is drastic enough to warrant being called partial suppletion. The examples in (21) only minimally represent the gamut of exceptional noun plurals that exist in modern Ket:

(21) A few examples of irregular stem changes triggered by plural suffixation

uu	'meadow'	*oo-ŋ*	'meadows'
qàj	'elk'	*qii-n*	'elk.PL'
sa'q	'squirrel'	*saa-n*	'squirrels'
qə'q	'corner'	*qə̄-n*	'corners'
dāp	'shoulder'	*daa-n*	'shoulders'
dīd [dī˙t]	'spruce grouse'	*dek-ŋ*	'spruce grouses'
taal	'otter'	*tak-ŋ*	'otters'

This brief sample suffices to illustrate that irregular plurals involve not only vowel and coda consonant changes, but sometimes an unexplained form of the nasal plu-

ral suffix, as well. The pair *sa'q* 'squirrel' → *saan* 'squirrels' involves the intervocalic loss of uvular /q/, which occurs elsewhere in Ket phonology; however, there are also plurals where this does not occur: *tə'q* 'finger' → *təqin* [tʌʁin] 'fingers'. In the case of *qə'q* 'corner' → *qə̄-n* 'corners', where /q/ does elide, the appearance of *-(V)n* rather than *-(V)ŋ* is unexpected since this noun belongs to the inanimate class. The choice of *-(V)ŋ* rather than *-(V)n* for animate class *taal* 'otter' → *tak-ŋ* 'otters' and *dīd* 'spruce grouse' → *dek-ŋ* 'spruce grouse.PL' is equally unexpected. A possible explanation is that the choice of suffix allomorph originally followed the general rules, and the surface appearance of velar /ŋ/ in the plurals of 'otter' and 'spruce grouse' as well as /n/ in 'corners' arose through some sort of formerly predictable phonological interaction. A full account of such diachronic changes would require a separate study.

In contrast to the relatively large number of cases of partial suppletion among Ket count noun singular and plural pairs, full suppletion is limited to the following two examples:

(22) Suppletive number pairs

ke'd	'person'	*de'ŋ*	'people'
ōks	'tree', 'pole'	*a'q*	'trees', 'poles'

The origin of these two full suppletive pairs is at present likewise unexplained.[9]

Some plural nouns do not show a suffix at all, but instead express plurality entirely through irregular tonal alternations, with or without concomitant vowel ablaut:

(23) Tonal or vowel ablaut plurals in Ket

kūl	'raven'	*kuul*	'ravens'
kɔ́qin	'fox'	*kəqín*	'foxes'
qópqun	'cuckoo'	*qopqún*	'cuckoo birds'
tīb	'dog'	*ta'b*	'dogs'
tōk	'axe'	*tòk* [tɔ̀ɣ]	'axes'
ēj	'tongue'	*èj*	'tongues'
sēs	'river'	*sàs*	'rivers'
hās	'shaman's drum'	*hàs*	'shaman's drums'

---

9 Porotova (1990: 48) and later Georg (2007: 100) suggested that *de'ŋ* 'people' represents a contraction of a regular plural *ked-eŋ*. Castrén (1858: 167) in fact did record the form *keädeŋ* as the plural of *ket ~ kiet* 'person', alongside *dʲeŋ* 'people'. However, the modern Ket word for person, which is pronounced [kɛ't], contains a coda that reflects Proto-Yeniseian *d, whereas the onset /d/ of *de'ŋ* is a reflex of PY *dʲ, a different phoneme. In any event, there are no other plausible cases of initial syllable collapse in Ket roots. The origin of this suppletive pair thus remains unexplained.

There is evidence that at least some of these suffixless exceptions arose when one of the canonical plural suffix forms *-(V)n* or *-(V)ŋ* was absorbed into the noun root (for reasons as yet unclear), similar to the historical development of English ablaut plurals like *mice* or *teeth*. The cognate plural of 'river' in some 18th century word lists representing extinct Yeniseian languages appears with the expected inanimate class suffix *-(V)ŋ*. This interpretation is also supported by the frequency of falling tone in irregular ablaut plurals, since the Ket falling tone [v̂] generally arose through the loss of a final stem element. Zero affixation or internal flection does not appear to have been an original technique of inflection in Yeniseian languages. Nevertheless, some of these irregular plural forms are shared across the family's primary branches and undoubtedly existed in Common Yeniseian, showing that irregularities in noun plural formation are deeply rooted in the family.

The few exceptional cases where the singular and plural noun form are identical can also be explained as resulting from the elision of a plural suffix. In these stems, however, suffix absorption did not cause tonal or vowel ablaut, again for reasons that are not clear:

(24) Number syncretism in a few Southern Ket nouns

bə'n	'duck'	*bə'n*	'ducks'
sūj	'mosquito'	*sūj*	'mosquitoes'
tūt	'black midge'	*tūt*	'black midges'

Grammatical number in this small set of nouns is covert, appearing only in the morphosyntax as part of verb agreement or possessive marking (compare the expression of number in English *this deer – these deer*). The other Ket dialects occasionally retain the original plural suffix in a few of these words, as when Central Ket *tutn* is compared to Southern Ket *tūt*, both of which are animate class plurals meaning 'black midges'. This suggests that number syncretism in such pairs arose from plural suffix erosion. Why the loss of a plural suffix triggered tonal or vowel changes in some stems (23) but not in others (24) requires further investigation of the language's historical phonology.

Though exceptions to almost every general rule of Ket noun plural formation abound, it is still possible to provide a cogent overall description. The list in (25) summarizes the various intersecting patterns and exceptions that are at play in assigning plural markers to Ket count nouns, beginning with lexically marked plurals and ending with the most general patterns.

(25) Summary of plural formation techniques in Ket count nouns
   a. A random assortment of lexically marked exceptions to the rules listed below, which involve suppletion, tonal or vowel ablaut, and number syncretism
   b. Stems ending in /n/ usually pluralize with *-(V)ŋ* regardless of meaning
   c. Inanimate class stems ending in /ŋ/ or /m/ and many ending in /s/ pluralize with *-(V)n*

d. Other inanimate class nouns and many grammatically animate class nouns denoting body parts, inanimate objects, or some lower forms of life pluralize with *-(V)ŋ*
e. Recent loanwords almost always pluralize with *-(V)ŋ* regardless of meaning
f. Most kinship terms pluralize with *-aŋ*, which typically attracts the pitch peak
g. Most other nouns denoting people or animals pluralize with *-(V)n*
h. *S*-nominalizations always pluralize with *-in* regardless of meaning or form

Irregularities in Ket noun plural formation offer a wealth of evidence for reconstructing the Proto-Yeniseian consonant inventory, which included more phonemes than modern Ket. Unstable final-stem /l/ that truncates in plural forms like *sèl* 'reindeer' → *se'-n* 'reindeer (plural)', and *sa'l* 'crucian fish' → *sa'n* 'crucian fish (plural)' can be reconstructed as a Proto-Yeniseian rhotic (or retroflex) sound rather than a lateral, as supported by its reflexes in Yugh *seʰr* 'reindeer' and *sa'r* 'crucian (fish)'. Stable /l/ in nouns like *būl* 'leg' → *bul-aŋ* 'legs' and *ool* 'container' → *ol-aŋ* 'covers' traces back to an original lateral, appearing also as /l/ in the Yugh cognates *būl* 'legs' and *o'l* 'cover'. Unstable coda /d/ in Ket, which participates in morphophonemic alternations with velar /k/ and corresponds to /dʲ/ in Yugh and /r/ in Kott, reflects the retroflex stop *dʳ in Proto-Yeniseian. One example is Kott (*feŋ-*)*čera* '(female) spruce grouse' and Ket *dīd*, Yugh *dīdʲ* 'spruce grouse', where the velar nasal in the Ket and Yugh irregular plural form *dekŋ* 'spruce grouse.PL' presumably arose because the alveolar nasal of the animate class suffix *-(V)*n* interacted with the retroflex stem coda.

   The foregoing discussion does not do justice to the full variety of irregular noun plurals in Yeniseian. The extensive coverage of dialectal forms in (Porotova 1990: 22–61), as well as the lucid presentations of plural allomorphy in Georg (2007: 92–101) and Werner (1997b: 96–102), consider many more exceptional plurals than could be analyzed here. The most complete reference for Southern Ket noun plurals is the two-volume *Comprehensive dictionary of Ket* (Kotorova & Nefedov 2015). Porotova's dialectal noun dictionary (Porotova 2002) is additionally useful because it includes plurals in Yugh as well as from all three Ket dialects.

2.3.2 Dual number marking and apparent pleonastic plural marking in Ket nouns

A diachronic perspective also explains how dual number marking arose in a few Ket nouns, and why other nouns appear to add two consecutive plural suffixes. Many Ket monosyllabic singular noun stems result from contractions of disyllabic root compounds or involve elided final-stem consonants. These changes sometimes have implications for understanding number marking in the modern language. Plural noun forms that seem to contain two consecutive plural suffixes include:

kū 'opening / mouth (of object)' → *kuniŋ* 'openings', *dɯ'* 'hat' → *dɯniŋ* 'hats', and *kɔ́jka* ~ *kəjá* 'head' → *kɔ́jkeniŋ* (alongside *kɔ́jken*) 'heads'. In the case of *kū* 'opening' and *dɯ'* 'hat', the stem originally ended in ⋆x, which elided in the bare singular form but was preserved through nasalization before the plural suffix *-(V)ŋ*. The forms *dɯ'* ~ *dɯn-* 'hat' and *kū* ~ *kun-* 'opening' are nothing more than irregular root allomorphs. The noun *kɔ́jka* 'head', on the other hand, is a compound of ⋆kɯj 'center / inner' + ⋆gen 'brain', so that *kɔ́jken* is the original singular stem, from which the final /n/ was later dropped in the singular form after being reanalyzed as a plural suffix. This yielded the truncated stem *kəjká* 'head', with *kɔ́jken* reanalyzed as a plural form alongside the original plural *kɔ́jkeniŋ* 'heads'. Other apparent examples of plural suffix concatenations listed by Porotova (1990: 155–158) or Werner (1997b: 96–97) can be explained in similar fashion.

In a similar way, dual marking developed in a few Ket nouns denoting naturally paired objects. In modern Ket usage the noun meaning 'sleeve' has a singular form *bánno*, a dual form *bánno-n*, and a plural (three or more) form *bánno-n-iŋ*. Etymologically, 'sleeve' is a compound of ⋆ben 'double' + ⋆gowx 'opening'. Its original structure is better preserved in the Yugh cognate *bengóu* 'sleeve'. The same process created *ókde* 'one ear', *ókde-n* 'pair of ears', *ógde-n-iŋ* '(three or more) ears', where the "dual" /n/ probably developed via nasalization of a fricative final-stem coda (analogous to ⋆gowx 'opening'). Another example is *ólta* 'one testicle', *ólta-ŋ* 'pair of testicles', *ólta-ɣ-in* '(three or more) testicles' (< *ol* 'covered' + *tə'ŋ* 'head'). This did not occur in *holtɔ́ŋ* 'button' → *holtɔ́kin* [hɔltʌɣin] 'buttons' (< *hol(ad)* 'leather' + *təŋ* 'head'), since buttons are not frequently paired objects. Only nouns that happened to denote naturally paired objects innovated the expression of dual number via semantic reanalysis of a nasal final coda as a dual marker. The rise of dual number in the Ket noun meaning 'eye', however, remains unexplained: *dēs* 'one eye', *dès* 'pair of eyes', and *destáŋ* 'three or more eyes'.

Nouns expressing naturally paired objects also developed a means to specially emphasize singular number by adding the modifier *qóleb* 'half': *qóleb dès* 'one single eye, one-eyed' (literally 'half eye'), *qóleb būl* 'one single leg (*būl*)', 'one-legged'. Georg (2007: 92) points out that this pattern, which is shared with Uralic, is likely a calque from one of the surrounding Siberian language families, where it is ancient and widespread.

A few remarks should be made here about the grammatical expression of number in determiner phrases consisting of an attributive numeral followed by quantified noun, as this topic is also tangentially relevant to the expression of dual number in Ket. Cardinal numeral words exist in two forms in Yeniseian. Forms used as attributive modifiers of nouns are normally unmarked for grammatical number. After *qo'k* '(animate class) one' and *qūs* '(inanimate class) one', nouns appear in their unmarked singular form: *qo'k ke'd* 'one person', *qūs qu's* 'one tent'. After *ūn* 'two', *dōŋ* 'three', *sīk* 'four', *qāk* 'five', and so forth, the plural form is used: *ūn de'ŋ* 'two people', *dōŋ qu'ŋ* 'three tents', etc. Porotova (1990: 171) observes, however, that

singular noun forms sometimes appear after *ūn* 'two', and still less often after *dōŋ* 'three' or *sīk* 'four': *ūn / dōŋ / sīk to'q* 'two / three / four steps'; after numbers five and higher, Ket nouns invariably appear in the expected plural form: *qāk / à / òn toq-ŋ …* 'five / six / seven steps', etc. Also, nouns that have developed a dual form show a tendency to use it after the numeral 'two', as in *ūn kə́jken* 'two heads' when compared to *qūs kə́jka* 'one head' and *qāk kə́jkeniŋ* 'five heads'. This usage does not appear to be fully regular, however, as *dōŋ kə́jken* 'three heads' (containing the so-called dual form of the noun) have also been recorded. Because heads are not a naturally paired object, perhaps the original stem *kə́jken* came to be treated as a sort of paucal form by certain individual speakers. It is also possible that interference from Russian, with its special treatment of using the genitive singular rather than the plural after the numbers two, three and four, may have influenced the rise of such forms.

Porotova (1990: 120) also lists a few rare examples where the nominalized form of the numeral is used attributively with a plural noun, leading to multi-site marking in the number phrase, including *un-aŋ de'ŋ* 'two people', with animate plural −*aŋ* on the numeral, in place of the expected *un deŋ* 'two people', where the numeral form is unmarked. Such cases do not appear to represent canonical usage. These rare examples also violate the typically strong head marking profile of Yeniseian morphosyntax. Only attributive forms of the numeral 'one' regularly and canonically reflect the grammatical class of the head noun and thus constitute genuine examples of dependent marking in modern Ket. See, however, Section 3.2 below for a description of two more widespread instances of marking on both modifier and noun in Ket noun and determiner phrases.

2.3.3 Other number-related aspects of Ket noun morphology

Pluralia tantum nouns are rare in Ket. One example is *kə't* 'children of the same mother'; the stem *dūl* 'child (in general)' is singular to the formally related forms *dulkat ∼ dulkitn* 'children (in general)', forming a pair that displays the typical singular/plural number contrast, like nearly all Ket nouns denoting countable objects. Two other possible examples are *obáŋ*, the plural of *ōb* 'father', and *amáŋ*, the plural of *ām* 'mother', when used in the secondary meaning of 'parents, ancestors'. Because the singular nouns *ōp* and *ām* can only mean 'father' or 'mother' and not '(generic) parent' or 'ancestor', the plural forms *obáŋ* and *amáŋ* when used to mean 'parents', 'ancestors' could be regarded as *pluralia tantum* nouns.

Mass nouns in Ket are often *singularia tantum* lacking plural forms, as is typical for many languages with a number distinction in their noun morphology: *ūl* 'water', *aal* 'broth', *altəq* 'mud', *ho'q* 'filth', *na'n* 'bread', *īm* 'pine nuts', *hu's* 'whortleberries', *tulit* 'red currants', etc. Many of these words derive count nouns by adding a suffix expressing a quantifiable amount of the given item. Such mass-quantifying

suffixes express specific amounts, shapes and consistencies. The most productive is the singulative suffix *-des ~ -dis*, which attaches to mass nouns to express a countable quantity in the form of a droplet of liquid or a single berry or pine nut. The examples in (26) juxtapose the underlying mass noun with its derived count noun and its plural form:

(26) Count nouns derived with the mass-quantifying suffix *-des ~ -dis*

ūl	'water', 'rain'	*úldis*	'droplet', 'raindrop'	(plural: *úldisn*)
īm	'pine nuts'	*ímdis*	'(single) pine nut'	(plural: *ímdisn*)
eel	'lingonberries'	*éldis*	'(single) lingonberry'	(plural: *éldisn*)
óŋniŋ	'roe'	*óŋndis*	'(single) fish egg'	(plural: *óŋndisn*)

The meaning of a single droplet or single discrete berry, nut, or grain of roe accrues from the suffix's etymological origin in the anatomical noun root *dēs* 'eye', which accounts for why the countable items it expresses are normally roundish is shape. Only rarely is this suffix used to convert a mass noun into a countable object of radically different shape, perhaps the sole example being *daan* 'grass' → *daandis* 'blade of grass (plural: *dandisn*).

There are a number of other, less productive mass-quantifying suffixes in Ket, all derived from roots or combinations of roots. The noun *lamtə ~ lamt* 'piece' can be added to Ket mass nouns to convey a congealed lump or broken off piece, thus deriving another set of countable noun stems:

(27) Count nouns derived by adding the suffix *-lamt* 'piece' to mass noun stems

kūd	'fat'	*kutlámt*	'lump of fat'	(plural: *kutlámtaŋ*)
sūl	'blood'	*sullámt*	'blood clot'	(plural: *sullámtaŋ*)
na'n	'bread'	*nanlámt*	'piece of bread'	(plural: *nanlámtaŋ*)
ōks	'wood'	*okslámt*	'(small) piece of wood'	(plural: *okslámtaŋ*)
qō	'ice'	*qoklámt*	'(small) chunk of ice'	(plural: *qoklámtaŋ*)

Yugh uses the suffix *-lap* in a similar way: *sur-láp* 'blood clot' (plural *sur-láf-ɯn*). The Ket noun *dápqul* 'heap', which originated via metathesis involving the roots *daq 'put / laid down' + *pɯl 'growth, mass, quantity', is added to a few mass nouns to denote countable larger pieces: *qokdápqul* '(large) chunk of ice' (plural: *qokdápqulaŋ*). Finally, the Ket suffix *-les*, of unknown etymology, derives count nouns expressing pieces of flat flexible objects: *si'k* 'rawhide' → *síkles* 'piece of rawhide' (plural *sikles-n*).

Some of these Yeniseian mass-quantifying suffixes have been called 'singulatives' (Porotova 1990: 65). Note, however, that they are derivational suffixes and cannot be regarded as grammatical markers of singular number because they remain when the stems they derive are inflected for plural meaning. In this connection, the count noun *óŋndis* '(single) fish egg', which is derived from *óŋniŋ* 'roe' is

noteworthy because it involves the truncation of the final-stem segments *-iŋ*, probably via reanalysis as a plural marker. The same process also seems to be at play in *hə́naŋdis ~ hə́ndis* 'grain of sand', where the final portion of the stem *hə́naŋ* 'sand' is optionally deleted, probably due to its resemblance to a plural marker. The mass noun *hə́naŋ* 'sand' also unusually allows pluralization as *hənaŋ-an* in the meaning 'expanses of sand', as recorded by Werner (2002, vol 1: 338). Still, the mass-quantifying 'singulative' suffix *-dis* in *hə́naŋdis ~ hə́ndis* 'grain of sand' cannot be viewed as a grammatical marker of singular number, since it is compatible with the addition of a plural suffix: *hə́naŋdis-n*

There is some evidence, however, that a genuine grammatical singulative suffix once operated in Yeniseian. Helimski (2016) identifies the /s/ element at the end of several noun stems as a fossilized singulative. This interpretation is supported by the fact that the element in question does disappear when a plural suffix is added.

(28) Possible examples of a fossilized singular marker (the "S-singulative") in Ket

tuʼs	'stone', 'rock'	*təʼŋ*	'stones', 'rocks'
quʼs	'birchbark tent', also 'house'	*quʼŋ*	'birchbark tents'
qeʼs	'sandbar (in river)'	*qedeŋ*	'sandbars'
ōks	'tree'	*aʼq*	'trees'

On the other hand, the disappearance of word-final /s/ in these forms might instead be interpreted as simply another irregular stem change in Yeniseian plural formation, with the elision of /s/ due to some as yet unexplained morphophonemic process. In particular, the forms for 'tree' are suppletive, which complicates any attempt to argue that /s/ in the singular form is a morpheme separable from the rest of the word. In any event, none of these forms can be used as mass nouns by deleting the final /s/; for instance, there is no bare form like [tuʼ] that means 'rock' as a generic, uncountable mass. The existence of an S-singulative in an earlier stage of Yeniseian, along with its possible etymological connection to the highly productive nominalizing suffix *-si*, remains plausible but unproven. Solving this problem would require additional diachronic investigation of irregular noun plural morphology.

Ket nouns normally have only one plural suffix, with apparent cases of consecutive nasal suffixes explainable as resulting from the loss of a final stem nasal from what was originally the full singular stem. However, genuine pleonastic (multi-site) marking of plurality does occur in some noun + noun compounds made through coalescence of two formerly independent word forms. A few examples, adapted from Porotova (1990: 158), are shown in (29), with hyphens inserted to show the morpheme breakdown:

(29) Noun compounds with multi-site number marking

búl-dal	'knee tendon'	< *būl* 'leg' + *daʼl* 'tendon'	→ plural:	*búl-aŋ-dál-eŋ*
ál-bes	'waist-length fur coat'	< *āl* 'half' + *beʼs* 'rabbit'	→ plural:	*ál-aŋ-bés-n*
sés-tas	'suede boot'	< *sàs* 'suede' + *tès* 'boot'	→ plural:	*sás-eŋ-tés-aŋ*

The first element in such compounds modifies the second to express a contiguity-based relationship involving parts and wholes. Their internal structure reflects the general syntactic pattern of modifier + head required for all Yeniseian noun and determiner phrases.

2.4 Verbal number

The polysynthetic Ket finite verb has a diverse arsenal of morphological means for expressing pluractionality. Only a few examples will be examined here.

Many stems use suppletion (full or partial) of the base (a term used by Ketologists to refer to the stem's rightmost lexical morpheme) in order to signal a contrast between a single event and multiple events affecting multiple entities separately:

(30) Base suppletion expressing types of pluractionality
 a. *d-aŋ-i-b-to*
    1SBJ-hanging-PRES-3INAN.OBJ-put.once
    'I hang it up.' 'I hang them up (all together, in one action)'

 b. *d-aŋ-i-b-uk*
    1SBJ-hanging-PRES-3INAN.OBJ-put.more.than.once
    'I hang them up (two or more separately, perhaps one after the other).'

 c. *t-a-b-qut*
    down-PRES-3INAN.SBJ-one.lies
    'It lies (there).'

 d. *t-a-b-damin*
    down-PRES-3INAN.SBJ-many.lie
    'They (many things) lie (there).'

The form (30b) *d-aŋ-i-b-uk* 'I hang them up' expresses multiple events, while the form (30a) *d-aŋ-i-b-to* 'I hang it up' expresses a single event and could also mean 'I hang them up' in cases where a handful of inanimate objects are simultaneously placed in a hanging position. The form (30d) *t-a-b-damin* 'They (objects) lie there' denotes a single situation, but with emphasis on more than one object lying in different places. This form is also interesting in that its composite base, *-damin*, seems to contain a fossilized plural suffix *-in*, which in all other types of finite verb forms correlates with animate class subjects only.

Along with base suppletion, the pluractional (PLTC) prefix *d-* sometimes appears in multiple-action stems.

(31) Example of (a) single- and (b) multiple-action pair involving pluractional *d-*

    a. *d-es-a-b-daq*
       1SBJ-surface-PRES-3ANIM.PL.OBJ-put.once
       'I lay it down.' / 'I lay them down (objects placed together in one action).'

    b. *da=d-a-b-da*
       3 F.SBJ=PLCT-PRES-3ANIM.PL.OBJ-put.many
       'I lay it down many times.' / 'I lay them down (objects, one after another).'

In another pair, the pluractional (PLCT) prefix *n-* distinguishes the multiple-action stem, in addition to base morpheme suppletion:

(32) a. *d-aŋ-baq*
       1SBJ-3ANIM.PL.OBJ-give.once
       'I give them (something) once.'

    b. *d-aŋ-n-bu*
       1SBJ-3ANIM.PL.OBJ-PLCT-give.many.times
       'I give them (something) many times.'

The pluractional prefix *n-* is unique to this pair, while *d-*, though also largely unproductive, is found in many stems, where it sometimes surfaces as *t-* through phonological merger with adjacent morphemes.

    Some verbs express pluractionality by using two pluractional affixes as well as a suppletion of the morpheme:

(33) a. *da=don-ba-h-ol-ted*
       3 F.SBJ=knife-1SG.obj-area-PST-hit.endwise
       'She stabbed me (once).'

    b. *da=don-an-ba-t-ol-do*
       3 F.SBJ=knife-PL-1SG.obj-PLCT.area-PST-gouge
       'She kept stabbing me.'

The pluractional suffix form /t/ in (33b) derives from *d-* merged with the area prefix *h-*, which remains visible in the single-action form in (33a). Notice also that the plural suffix *-an* on the incorporated instrument noun *do'n* 'knife' is a pluractional marker in that it expresses multiples acts of stabbing (done by one person or by many people) rather than stabbing with multiple knives. The noun plural suffix in this mildly productive verb stem pattern therefore marks pluractionality in conjunction with the addition of the pluractional prefix *d-* as well as base morpheme suppletion – though both *-ted* 'hit endwise' and *-do* 'cut, gouge, hew' when used on their own can refer to either single or multiple actions.

Just as multi-site grammatical agreement is not infrequent in Ket, the lexical category of pluractionality is also frequently expressed using combinations of multiple morphemes in the same verb form (31–33). There is no evidence that Ket pluractionality involves a morphological distinction between a small number events (two or three) and a larger number.

3 Agreement and syntax of number

The marking of grammatical categories on modifiers is generally atypical for Yeniseian, being limited to a few types of noun or determiner phrases. As mentioned in 2.2.3, attributive forms of the numeral 'one' regularly agree with their head noun in animacy but not gender: animate class *qo'k qīm* 'one woman' and *qo'k hīk* 'one man' vs. inanimate class *qūs tu's* 'one rock'. Marking of plural number on both the numeral form and the noun has been documented sporadically in a few phrases with higher numbers like *un-aŋ de'ŋ* 'two people' in place of the canonical *ūn de'ŋ* 'two people', which shows the typical pattern of plural marking on the noun only. This section examines two additional instances where number is marked on modifiers as well as on the noun. In the first instance, examined in 3.1 below, demonstrative pronouns are grammatically marked for noun class (both animacy and gender) as well as animate plural number, while the nouns they modify are marked only for plural number, but not for class. In the second, a small minority of adjectives and action nominals (the Ket equivalent of participles) take a plural suffix that distinguishes them from their singular forms (3.2). There are no other instances in Ket nominal morphology where modifiers take grammatical affixes. Case suffixes in noun or determiner phrases attach to the head noun only and never appear on modifying words. Adjectives and adverbs that undergo *S*-nominalization can take case inflections, but only as the head of a noun phrase. Finally, subsection 3.3 examines number marking in the subject complements of locative and existential clauses.

3.1 Class and number agreement in demonstrative pronouns

The most frequently encountered and straightforward exception to the general pattern of Ket head marking appears in determiner phrases, where demonstrative pronouns agree in class and number with their head noun. Ket demonstrative roots express a three-way contrast between proximal *ki-* (close to speaker), medial *tu-* (farther from speaker but earlier mentioned or in the general vicinity) and distal *qa-* (distant from speech situation). When modifying a singular masculine animate class noun, all three demonstrative roots take the suffix *-d* (pronounced as the rhotic flap [ɾ] in southern Ket). When modifying a feminine singular animate class or inanimate

Tab. 3: Class and number agreement forms of demonstrative pronouns.

-d *kīd / tūd / qād hīk* 'this / that (near) / that (far) man'	*-ne* *kine / tune / qane hikn* 'these / those (near) / those (far) men'
-de *kide / tude / qade qīm* 'this / that (near) / that (far) woman'	*kine / tune / qane qimn* 'these / those (near) / those (far) women'
kide / tude / qade qu's 'this / that (near) / that (far) tent'	*-de* *kide / tude / qade qu'ŋ* 'these / those (near) / those (far) tents'

class noun of either number, they take the suffix *-de* (pronounced [rɛ] in Southern Ket). When modifying any plural animate class noun, they take the suffix *-ne*.

Plural marking in demonstratives that modify animate class nouns represents another case where grammatical marking appears on the modifier as well as its head noun. Inflection for class (masculine animate, feminine animate, and inanimate) in singular forms of demonstrative modifiers constitutes dependent marking, however, since these categories are covert in the head nouns themselves.

Vajda (2013a: 87–88) argued that the agreement suffixes in modern Ket demonstrative pronouns are vestiges of ancient possessive morphology, which explains why they exactly mirror the grammatical categories expressed by 3rd person possessive clitics.

3.2 Plural marking in adjective and action nominal forms

While most Ket adjectives keep the same form regardless of whether they apply to a single or plural entity, about a dozen mark plurality by adding the suffix *-(V)ŋ* (Bibikova 1976: 91). Although this suffix, with its epenthetic vowel and velar nasal, resembles that used on inanimate class plural nouns, these exceptional adjective plural forms modify inanimate as well as animate class nouns, as can be seen in these attributive phrases. Example (34) provides an exhaustive list of Ket adjective forms that typically occur in different singular and plural forms:

(34) Examples of plural agreement in adjectives

qà ke't	'big person'	*qē-ŋ de'ŋ*	'big people'
ugd būl	'long leg'	*ugd-eŋ bulaŋ*	'long legs'
ho'l qa'd	'short coat'	*hol-aŋ qadaŋ*	'short coats'
bo'l dᴜlkit	'fat child'	*bol-aŋ dᴜlkitn*	'fat children'
ka't ke'd	'old person'	*kat-aŋ de'ŋ*	'old people' [human age]
sīn ōks	'old / rotten tree'	*sin-aŋ a'q*	'old / rotten trees' [objects]

dɔ́qta qūut	'fast wolf'	*dɔ́qta-ŋ qǝtn*	'fast wolves'
uul la'm	'smooth board'	*uul-aŋ lemiŋ*	'smooth boards'
ǝǝl hīk	'unmarried man'	*ǝǝl-aŋ hikn*	'unmarried men'
èt ko'p	'living chipmunk'	*et-iŋ koon*	'living chipmunks'
ēt ìt	'sharp tooth'	*et-iŋ iteŋ*	'sharp teeth'
būud àŋ	'strong rope'	*bud-eŋ aŋen*	'strong ropes'
to't de'	'shallow lake'	*tot-iŋ dēŋ*	'shallow lakes'

Using internal reconstruction, Vajda (2013b: 20–22) demonstrated that the element *-(V)ŋ* in these words was originally an adjective-deriving suffix that was reanalyzed as a marker of plurality through false analogy with the common noun plural suffix *-(V)ŋ*. For reasons that are not entirely clear, some adjectives absorbed the suffix into their roots in plural as well as singular forms, which explains the unusually high prevalence of falling tone in monosyllabic Ket adjectives. Other adjectives retain the suffix in all contexts, suggesting it was originally part of the stem rather than a plural marker. Two examples of the latter type of adjective stems are Ket *sokŋ* 'thick', *údokŋ* 'lazy', which modify either singular or plural nouns: *údokŋ ke't* 'lazy person', *údokŋ de'ŋ* 'lazy people'. Other adjectives show free variation between forms with or without the suffix. One example is Yugh *súrbes ~ surbèːʰs ~ súrbesiŋ* 'red' < **sūr* 'blood' + **wes* 'resemble' + **(V)ŋ* 'ADJ suffix', with any one of these variants capable of modifying either singular or plural nouns. The Ket cognate *súlem* 'red', which shows a more radical reduction, also occurs with either singular or plural nouns – *súlem tɯ's* 'red stone', *súlem tǝ'ŋ* 'red stones', providing yet another indication that the final nasal element was originally part of the stem. Recognizing this semantic reanalysis of the adjective derivational suffix **-(V)ŋ*, which appears to be an innovation that is limited to the Ket-Yugh branch of the family, makes it unnecessary to posit adjective number agreement as an original feature of Yeniseian morphology and also explains the origin of this highly idiosyncratic trait in a strongly head marking language like Ket.

An identical process affected the stem phonology of action nominals, a lexical class of modifiers that serves a participial as well as infinitival function in Ket. Action nominal derivation originally involved attaching a nasal suffix to a finite verb root or complex lexical stem. This suffix is homonymous with the adjective-deriving suffix *-(V)ŋ* and may be etymologically identical to it; in any event, both adjectival and action nominal *-(V)ŋ* underwent a parallel evolution in the historical development of Ket. In some action nominals this suffix remained unchanged in all forms: *bágdeŋ* 'pulling', *táseŋ* 'getting up', *ániŋ* 'playing', *qódeŋ* 'dying', *íliŋ* 'eating', *éjiŋ* 'going', *háleŋ* 'wrapping up', etc. In other stems it was absorbed into the root, usually creating falling tone: *bɔ̀k* [bʌ̀ɣ] 'finding', *tàd* [tàr] 'beating', *tòs* 'raising', *kàl* 'fighting', *dàq* [dàʁ] 'laughing', *dùd* [dùr] 'burning', *èj* 'killing', etc. Less often, absorption of the suffix into the verb root yields another tone: *kīt* 'rubbing', *ān* 'cooking', *kǝ'j* 'going hunting', *dǝ'q* 'living', etc. A full explanation for why the action nominal and

adjective suffix was sometimes retained and sometimes absorbed into the stem would require a detailed excursion into diachronic phonology, similar to what would be required to explain the origin of exceptional ablaut noun plurals. Parallelling the development of a number contrast in a few adjective stems, this suffix was similarly reanalyzed as a marker of plurality or pluractionality in a minority of action nominal stems. This has resulted in a sporadic contrast between a contracted form used with reference to single actions or singular entities, and a suffix-augmented form that refers to multiple actions or modifies plural nouns. Example (35) shows three Ket action nominals that typically use distinct singular and plural forms.

(35) Number marking in action nominals
suulbèd 'sled-making (one event)'   *suulbed-eŋ* 'sled-making (in general)'
təəl i' 'freezing day'   *təl-iŋ ekŋ* 'freezing days'
ūs i' 'warm day'[10]   *us-eŋ ekŋ* 'warm days'

The same subset of adjective and action nominal stems that show a number contrast when used as modifiers in noun phrases carries this opposition into the stem morphology of nominalizations. Example (36) shows compound words containing adjective suffixes that have been reanalyzed as plural markers:

(36) a.

noun phrase	compound adjective	*S*-nominalization	plural of *S*-nominalization
bol-eŋ lon-eŋ	*bol-eŋ-lon-eŋ ke'd*	*bol-eŋ-lon-eŋ-s*	*bol-eŋ-lon-eŋ-s-in*
thick-PL lip-PL	thick-PL-lip-PL person	thick-PL-lip-PL-NMLZ	thick-PL-lip-PL-NMLZ-PL
'thick lips'	'thick-lipped person'	'ones with thick lips'	'ones with thick lips'

b.

noun phrase	compound adjective	*S*-nominalization	plural of *S*-nominalization
ukd-eŋ bul-aŋ	*ukd-eŋ-bul-aŋ assel*	*ukd-eŋ-bul-aŋ-s*	*ukd-eŋ-bul-aŋ-s-in*
long-PL leg-PL	long-PL-leg-PL animal	long-PL-leg-PL-NMLZ	long-PL-leg-PL-NMLZ-PL
'long legs'	'long-legged animal'	'one with long-legs'	'ones with long legs'

In complex nouns or noun phrases containing more than one plural marker, it is the number category at the noun's rightmost edge that signals whether a singular or plural entity is being referred to. The retention of this type of multi-site number marking in complex stems, however, is highly lexicalized, as can be seen from the forms in (37):

(37)

singular phrase	plural phrase	*S*-nominalization	plural of *S*-nominalization
ukd a'd	*ukd-eŋ ad-eŋ*	*ukd-ad-s*	*ukd-ad-s-in*
long bone	long-PL bone-PL	long-bone-NMLZ	long-bone-NMLZ-PL
long bone	'long bones'	'one with big bones'	'ones with big bones'

---

10 Though unclear from the English translation 'warm day', the morpheme *us* is a verb root that heads finite forms like *b-il-us* 'it turned warm', in addition to appearing in the action nominal *ūs ~ useŋ* 'turning / being warm'.

Here the use of plural marking in both modifier and noun fails to carry through to the complex stem form nominalized by -s, in contrast to the presence of precisely two lips or precisely two (or four) legs. The NP 'long bone' therefore acts more like a mass noun in the complex nominalizations shown in (37). At the same time, note the lack of plural adjective form in *ukde-həŋn-eŋ ke'd* 'long-armed man', a compound recorded by Werner (2002, vol. 2: 322). This type of variation probably follows from the originally derivational nature of the adjective suffix *-(V)ŋ*, whose grammatical function of marking plural agreement developed only later through semantic reanalysis and remained sporadic.

This secondary singular/plural distinction in the same subset of adjective and action nominal stems also shows up when these forms are incorporated into finite verbs (3.4.1) or are used as subject complements in copular clauses (3.3).

3.3 Number marking in the subject complements of linking verb clauses

While most Ket clauses are headed by a finite verb, some copular clauses contain a subject complement and no tense-bearing verb form at all. The predicates in clauses of this type have three different kinds of subject complements: 1) a bare noun or noun phrase; 2) an *S*-nominalization derived from another part of speech; or 3) a qualitative adjective or locational stem inflected with a subject agreement suffix. Bare nouns used as subject complements may be singular or plural in form but lack agreement in person with their clausal subject:[11]

(38) Nouns used predicatively as subject complements

    a. *tu-de   tu's*
       that-INAN rock
       'That is a rock'

    b. *tu-de   tə'-ŋ*
       that-INAN rocks-PL
       'Those are rocks.'

    c. *tū-d   qòj ben ke'd*
       that-M bear really person
       'That bear is really human.'

---

11 During his 19th century fieldwork with Ket and Yugh speakers, Castrén (1858: 100–103) recorded noun forms like *uob-di* 'I am a father' and *uob-du* 'He is a father' with subject concord suffixes, but such forms are no longer used in any of the three modern Ket dialects (Werner 1997b: 306).

d. *ətn de'ŋ*
we people
'We are people.'

The second morphological type of subject complement in copular clauses with no finite verb form is the *S*-nominalization, which can be derived from any non-noun. Used predicatively, *S*-nominalizations regularly show number agreement with the clause subject; however, similar to bare nouns they lack person agreement. Adjectives like *qà ~ qēŋ* (or *qàŋ*) 'big' that display a morphological number contrast when used attributively, maintain this formal contrast when nominalized by *-s*:

(39) S-nominalizations used predicatively as subject complements in copular clauses

a. *tu's qà-s*
rock big-NMLZ
'The rock is big.'

b. *tu'-ŋ   qe-ŋ-s-in*
rock-PL big-PL-NMLZ-PL
'The rocks are big.'

c. *ke'd   qà-s*
person big-NMLZ
'The person is big.'

d. *de'ŋ   qe-ŋ-s-in*
people big-PL-NMLZ-PL
'The people are big.'

S-nominalizations show plural agreement with inanimate as well as animate class subjects. In addition, forms like *qe-ŋ-s-in* 'they (either animate or inanimate class) are big' display multi-site plural marking, while *S*-nominalizations derived from non-variable stem adjectives (the majority of forms) mark plurality word-finally only: *ke'd udokŋ-s* 'the person is lazy', *de'ŋ udokŋ-s-in* 'the people are lazy', *at aqta-s* 'I am good', *ətn aqta-s-in* 'we are good'.

The third type of non-finite predicate contains an adjective or locational adverb that takes a suffix agreeing in person, number, and class (animacy and gender) with the clause subject, as shown in (44):

(40) Concord suffixes on predicate adjectives used as subject complements

a. *tu's qa-am* [qaɣam]
rock big-INAN.SBJ
'The rock is big.'

b. *tə'-ŋ   qe-ŋ-am*
rock-PL big-PL-INAN.SBJ
'The rocks are big.'

c. *at qa-di*
I   big-1SG.sbj
'I am big.'

d. *ətn qe-ŋ-dəŋ*
we   big-PL-1PL.SBJ
'We are big.'

In contrast to S-nominalizations, there is no singular/plural differentiation in the inanimate class forms, except in the minority of stems where adjectival *-ŋ* has been reanalyzed as a plural marker. The concord suffixes that appear on predicate adjectives, locational adverbs, or action nominals[12] used as subject complements mark person and class, in addition to number. Just as in other realms of Yeniseian morphology, the masculine and feminine animate singular classes merge into a single animate plural class.

The examples in Table 4 show the full array of predicate concord suffixes on Ket and Yugh forms of the adverb **qapə*, which originally meant something like 'inside one's tent'.[13] The Yugh forms are adapted from Werner (1997a: 209).

Vajda (2019: 68) argued that modifying words with predicate concord suffixes were originally finite verbs constructed from an incorporated modifier followed by a

Tab. 4: Paradigm comparing Ket and Yugh predicate concord suffix forms.

	Southern Ket	Yugh
1SG 'I am at home / in my tent.'	*qaˑ-di*	*χaˑb-di'*
2SG 'You are …'	*qaˑ-ku*	*χaˑp-ku'*
3M.SG 'He is …'	*qaˑ-du*	*χaˑb-du'*
3 F.SG 'She is …'	*qaˑ-də*	*χaˑb-da'*
1PL 'We are …'	*qaˑ-dəŋ*	*χaˑb-dəːʰŋ*
2PL 'You (all) are …'	*qaˑ-kəŋ*	*χaˑp-kəːʰŋ*
3ANIM.PL 'They (people, animals) are …'	*qaɣ-aŋ*	*χaf-eːʰŋ*
3INAN 'It is …' 'They (things) are …'	*qaɣ-am*	*χaf-e'*

12 Subject concord suffixes on action nominals express that the subject is able to perform the given action': *il-di* 'I can sing' (from *i'l* 'singing', 'to sing').
13 The phonemic form of this adverb in modern Ket is /qap/, which could also be transcribed as /qah/, since word-final /p/ and word-initial /h/ represent a single phoneme in the transcription used in this article. The forms in (45) show the surface allomorphs [qaˑ] and [qaɣ] rather than the underlying phonemic form.

subject agreement prefix on an archaic word-final verb root meaning 'be' or 'having become', probably shaped something like *eŋ. Phonological reductions in this agreement prefix + verb root sequence produced the modern concord suffix forms, in which the original verb root was reanalyzed as a nasal plural agreement suffix. This interpretation also explains why predicate concord in modern Ket is suffixing, whereas subject and object agreement in finite verbs (described in the next section) follows a mostly prefixing template. If subject complements with predicate concord do indeed derive etymologically from prefixed finite verb forms, their adjective or locational adverb portions were once incorporated modifiers exactly like those still found in many finite verb forms today.

3.4 Number marking in finite verb forms

Ket verb morphology is based on a discontinuous stem that allows the incorporation of certain types of modifying elements (Vajda 2017a). Agreement affixes interdigitated between the stem's lexical morphemes index the subject and direct object arguments. Other affixes in the verb complex distinguish past and non-past tense. Finite verb clauses in Ket are exclusively head marking. While subjects and direct objects are indexed verb-internally, any noun or pronoun arguments appearing in the finite verb clause remain unmarked for grammatical relations. Yeniseian nominal morphology lacks any type of nominative, accusative, absolutive or ergative case markers altogether. As earlier described with reference to possessive clitics and other morphologies evolved from them, plural number agreement in finite verbs is grammatically marked only with reference to animate class subjects or objects. Inanimate class subjects or objects trigger the same affixal form regardless of whether they are singular or plural. The discussion below skirts around much of the verb's internal polysynthetic complexity to concentrate on features relevant to the topic of number marking. Subsection 3.4.1 examines how the finite verb template expresses grammatical subject and object number agreement. The main point is that plural agreement is normally limited to animate class subjects and objects. Subsection 3.4.2 draws on the description given earlier in 3.2 of how number marking in adjectives and action nominals arose via reanalysis in order to examine features of verb morphology originally unconnected with number that later acquired plural or pluractional meaning based on coincidental homonymy with the common noun plural suffix *-(V)ŋ*. Plural markers that arose in Ket via reanalysis of derivational or aspectual affixes differ from the inherited Yeniseian agreement system in occasionally reflecting the plurality of inanimate class subjects.

3.4.1 Verb-internal class, person, and number concord

Subject/object concord in finite verb forms is accomplished using verb-internal affixes that express the same grammatical categories of agreement in person, class

and number found in possessive constructions. Intransitive verb forms normally require subject agreement, while transitive verbs agree with their subject and direct object. The following template underlies every Ket finite verb form:

Tab. 5: Ket finite verb template, with agreement slots shaded.[14]

8	7	6	5	4 / 3	2	1	0	-1
SBJ person	incorporated noun, modifier, or action nominal	OBJ or SBJ	thematic consonant prefix(es)	3 SBJ or OBJ (originally in slot 4) now fused with conjugation marker (*s/i ~ a/o*) (originally in slot 3)	tense-mood aspect (*n ~ l*)	SBJ or OBJ	BASE (verb root or lexical aspect marker)	ANIM PL SBJ suffix

Each stem lexically chooses which of the shaded slots must be filled to express subject or object agreement. No verb form fills all of the template's five (shaded) agreement slots, though multi-site subject marking involving two or even three of these slots filled simultaneously in single verb forms is not uncommon. There are five productive intransitive agreement configurations and three productive transitive configurations, in addition to a smattering of unproductive types – a system first explained in Vajda (2001) and Vajda (2004: 44–76) and now most succinctly described in Nefedov & Vajda (2015: 38–48). Vajda (2017b) explains in detail the diachronic processes that gave rise to this typologically unusual system. Table 6 shows basic allomorphs found in each agreement slot.[15]

Of the positions left blank in Table 6, slots 7, 5, and 0 contain lexical morphemes, while slot 2 is used to mark tense, mood and aspect. Slots 6, 4, and 1 are used variously to express class, person, and number concord. By contrast, the word-initial marker in slot 8 expresses person and class agreement only, since it works in tandem with the template's only productive suffix, located word-finally, which expresses plurality of any animate class subject in verbs that use prefixal slot 8 for subject person and class agreement:

---

14 Subject complements with predicate concord suffixes of the type examined in 3.3 above, are composed of morphemes that once occupied slots 7-1-0, which reflects an intransitive configuration still productive today for several other kinds of finite verb stems.

15 See Nefedov & Vajda (2015: 38–40) for more detail on agreement marker allomorphy, including rules for when P8 markers surface as prefixes and when they surface as clitics.

Tab. 6: Agreement marking series in the Ket finite verb template.

8	7	6	5	4 / 3	2	1	0	−1
SBJ *di-* 1person *ku-* 2person *du-* 3M.SG *də-* 3F.SG		OBJ or SBJ *ba- ~ bo-* 1SG *ku-* 2SG *a- ~ o-* 3M.SG *i- ~ u-* 3F.SG *Ø- ~ u-* 3INAN *dəŋ-* 1PL *ku-* 2PL *aŋ- ~ oŋ-* 3ANIM.PL		3 SBJ or OBJ *(d)a-* 3M.SG[16] *(d)i-* 3F.SG *(d)aŋ-* 3ANIM.PL *b-* 3INAN		SBJ or OBJ *di-* 1SG *ku-* 2SG *ə-* 3ANIM.SG or 3INAN *dəŋ-* 1PL *kəŋ-* 2PL *əŋ-* 3ANIM.PL		ANIM PL SBJ suffix *-in*

(41) a. *di-k-a-qut* [diɣaʁut]
    1SBJ-up-PRES-walk
    'I ascend / go upward.'

   b. *di-k-a-qut-n*
    1SBJ-up-PRES-walk-ANIM.PL.SBJ
    'We ascend / go upward.'

   c. *ku-k-a-qut*
    2SBJ-up-PRES-walk
    'You.SG ascend / go upward.'

   d. *ku-k-a-qut-n*
    2SBJ-up-PRES-walk-ANIM.PL.SBJ
    'You.PL ascend / go upward.'

   e. *du-k-a-qut*
    3SBJ-up-PRES-walk
    'He ascends / goes upward.'

   f. *du-k-a-qut-n*
    3SBJ-up-PRES-walk-ANIM.PL.SBJ
    'They (animate, either gender) ascend / go upward.'

   g. *də-k-a-qut*
    3F.SG.SBJ-up-PRES-walk
    'She ascends / goes upward.'

---

16 The prefixes *d-* (ANIM) and *b-* (INAN) once filled slot 4. Today *b-* fills slot 3, having metathesized rightward ahead of the *a-*conjugation marker. The innovative Ket/Yugh gender contrast between MASC.SG *a-* and FEM.SG *i-* has eclipsed generic ANIM *d-*, which survives vestigially in only a small minority of forms.

In this common intransitive agreement pattern, the marker P8 *də- ∼ da=* indexes 3rd person singular feminine animate class subjects.[17] It is never used to index animate plural subjects, even those specifically denoting groups of females. Instead, the prefix *du-* appears in verb forms with animate plural subjects. The position 8 marker *du-* also indexes a 3rd person masculine animate subject.

The rest of this subsection examines number marking differences between animate and inanimate class arguments, as well as several cases of multi-site animate plural agreement.

Animacy plays just a pervasive role in Ket finite verb as elsewhere in Ket morphology. Despite this fact, Yeniseian lacks true semantic (active) alignment: roughly synonymous verbs often require different agreement configurations, while active and inactive intransitives with completely different semantics often follow an identical agreement pattern. Animate singular nouns display a formal dichotomy between masculine and feminine subclass, as has already been observed in predicate concord suffixes and other areas of the morphosyntax. Inanimate class subjects and objects, on the other hand, require a mostly different set of agreement markers that lack any number contrast between singular and plural. Finally, it is worth mentioning that inanimate class entities rarely serve as the subjects of transitive verbs in Ket. The sentences in (42) show examples of verb-internal subject concord in clauses containing arguments of different noun classes and grammatical numbers. The verb stem in these examples is a linking verb that incorporates its subject complement and uses slot 6 to express subject agreement:

(42) Intransitive verbs showing 3rd person animate class number agreement

    a. *qīm   baam-i-tonoq*
       woman old.woman-3 F.SBJ-became
       'The woman became (very) old.'

    b. *hīk   baat-a-tonoq*
       man old.man-3M.SBJ-became
       'The man became old.'

    c. *qim-n   baam-aŋ-aŋ-tonoq*
       woman-PL old.woman-PL-3ANIM.PL.SBJ-became
       'The women became old.'

    d. *hik-n   baat-aŋ-aŋ-tonoq*
       man-PL old.man-PL-3ANIM.PL.SBJ-became
       'The men became old.'

---

17 The P8 marker *də- ∼ da=* is also used to index an inanimate class subject in the few verbs of this agreement configuration that allow inanimate subjects, such as *qām hīk da=kasonam* 'an arrow took [= killed] the man'.

Just as in possessive morphology, the masculine and feminine animate distinction disappears in the plural: the same animate class plural affix *-aŋ* is used for plural nouns that denote males as well as females. This set of forms also illustrates how nouns that fulfill the role of subject complement retain their logical expression of number, which in plural subject forms like (42c) and (42d) produces yet another example of multi-site plural marking in Ket.

Inanimate class subjects and objects, by contrast, show no grammatical number distinction, as shown in (43d) which sets them apart from the rich array of animate class object agreement markers in (43a), (43b), and (43c):

(43) Class and number marking of objects in one set of Ket transitive verb forms
   a. *ku-a-ted* [kuɣatɛt]
      2SBJ-3M.OBJ-hit
      'You hit him.'

   b. *ku-i-ted* [kuˑtɛt]
      2SBJ-3 F.OBJ-hit
      'You hit her.'

   c. *k-aŋ-a-ted*
      2SBJ-3ANIM.PL.OBJ-PRES-hit
      'You hit them (people or animals).'

   d. *ku-b-ted*
      2SBJ-3INAN.OBJ-beat
      'You hit it.' / 'You hit them (inanimate objects).'

   e. *k-aŋ-a-teɣ-in*[18]
      2SBJ-3ANIM.PL.OBJ-PRES-hit-ANIM.PL.SBJ
      'You (plural) hit them (people or animals).'

   f. *ku-b-teɣ-in*
      2SBJ-3INAN.OBJ-beat-ANIM.PL.SBJ
      'You (plural) hit it.' / 'You (plural) hit them (inanimate objects).'

Other verbs require multi-site subject marking, with both person and number agreement marked twice in different slots in the verb form:

(44) a. Example of multi-site subject marking in an intransitive verb
      d**-aka-**dəŋ**-t-l-aq-**in
      1SBJ-upland-**1PL.SBJ**-to-PST-go-**ANIM.PL.SBJ**
      'We made a quick round trip from river to forest.'

---

18 The verb base *-ted* is replaced by the allomorph *-tek* whenever the animate class plural subject agreement suffix *-in*, a combination that is pronounced [tɛɣin].

b. Example of multi-site subject marking in a transitive verb
d-*əla*-**dəŋ**-*a*-*qos*-**n**
1SBJ-outside-**1PL.SBJ**-3M.OBJ-take-**ANIM.PL.SBJ**
'We take him outside.'

Several reasons for the origin of multi-site agreement markers in Ket verb stems have been identified (Vajda 2017a), including semantic reanalysis of an incorporate element that happens phonologically to resemble an agreement marker, but the origin of the phenomenon in examples (44a) and (44b) remains unclear.

All of the example verbs given so far are single-action stems. The next subsection addresses how the Ket verb expresses pluractionality.

3.4.2 Number marking in finite verbs arising through reanalysis

The number distinction developed in some adjectives and action nominals via reanalysis of the derivational suffix *-(V)ŋ* as a plural marker, as earlier described in 3.2, is maintained when these words are incorporated into finite verb forms. In verb forms with inanimate class subjects or objects, which do not mark plural agreement grammatically, the plural marker in the incorporated adjective is the sole marker of plurality, as in (45b) and (46b):

(45) Plural marking of incorporated adjective that correlates with inanimate class subject
a. *qa-d-a-b-qan*
big-TRANSITION-PRES-3INAN.SBJ-become
'It becomes big.'

b. *qe-ŋ-d-a-b-qan*
big-PL-TRANSITION-PRES-3INAN.SBJ-become
'They (inanimate objects) become big.'

(46) Plural marking of incorporated adjective that correlates with an inanimate class object
a. *d-ukde-t-a-p-sin*
1SBJ-long-cause-PRES-3INAN.OBJ-be
'I lengthen it.'

b. *d-ukde-ŋ-t-a-p-sin*
1SBJ-long-PL-cause-PRES-3INAN.OBJ-be
'I lengthen them (inanimate objects).'

In verbs where such adjectives correlate instead with animate class subjects or objects, which do regularly express grammatical plural agreement, the extra plural

suffix in the incorporated adjective form results in multiple exponence of plural marking, sometimes in tandem with other stem changes like base suppletion that likewise express plurality of the argument or pluractionality, as in (47b).

(47) Plural marking of incorporated adjective that correlates with an animate class subject
    a. *t=qa-d-a-qan*
       3M.SBJ=big-TRANSITION-PRES-become
       'He becomes big / grows up / matures.'

    b. *t=qe-ŋ-d-a-set-n*
       3ANIM.SBJ=big-PL-TRANSITION-PRES-many.become-ANIM.PL.SBJ
       'They (people or animals) become big / grow up / mature.'

Generally, adjectives that lack a number contrast outside the verb do not vary to express plurality when incorporated into a finite verb stem either.

(48) Examples of incorporated adjectives that lack a number distinction
    a. *d-aqta-a-qan* [daqtaʁan]
       3M.SBJ-good-PRES-become
       'He recovers.'

    b. *d-aqta-a-set-n*
       3ANIM.SBJ-good-PRES-many.animate.become-ANIM.PL.SBJ
       'They (people or animals) recover.'

Like variable adjective stems, variable forms of action nominals incorporated into finite verbs also signal a number contrast reflecting an intransitive subject or transitive object. Action nominal forms that retain the suffix *-(V)ŋ* tend to appear in multiple action verbs and contribute to the expression of the idea of pluractionality (PLCT), while this element usually is absent in verbs expressing single events:

(49) Variable action nominal forms used to distinguish single from multiple events
    a. *d-ulad-q-a-it*
       1SBJ-petting-into-3M.OBJ.PRES-put.once
       'I pet / start petting him (male dog, once).'

    b. *d-ulad-**iŋ**-q-a-da*
       1SBJ-petting-**PL**-into-3INAN.OBJ.PRES-put.many.times
       'I pet him (male dog, on many different occasions).'

The use of the reanalyzed action nominal suffix *-(V)ŋ* as a marker of pluractionality is fairly sporadic. It sometimes remains in Southern Ket single-event verbs as well, and there is even more variability across the Ket dialects. For example, Central Ket

d-uladiŋ-q-a-it 'I pet him once' and Yugh *d-ulʲaduɯŋ-χ-âj-it* 'I pet him once' show that the suffix is retained in single-event verb forms.

One final instance of a morpheme reanalyzed as a plural (pluractional) marker through coincidental homonymy with the common plural marking *-(V)ŋ* involves the archaic change-of-state (inchoative) suffix *-(V)ŋ*. In a few dozen stems, this old aspect suffix survived only in forms with plural subjects, where it was reinterpreted as a subject plural marker. With the base *-teel* 'freeze', it is sporadically retained when referring to plural entities, including inanimate class plurals.

(50) *b-in-teliŋ*
      3INAN.SBJ-PST-freeze
      (also recorded as *b-in-teel*, with loss of suffix causing root vowel lengthening)
      'They (inanimate class things) froze (once).'

Example (50) shows that plural markers arising through reanalysis can reference inanimate as well as animate class entities, while canonical plural agreement in finite verb forms only indexes animate class subjects or objects. Most other stems where it survives, however, are generally restricted to animate class subjects, as is the case with the base *-doq* 'fly'.

(51) Animate plural marking expressed by a reanalyzed inchoative suffix
      a. *d-in-doq*
         3ANIM.SBJ-PST-flew
         'He flew.'

      b. *d-in-aŋ-doq-**ŋ***
         3ANIM.SBJ-PST-3ANIM.PL.SBJ-flew-**PL**
         'They flew (together as a group).'

      c. *d=t-ol-aŋ-doq-**ŋ***
          3ANIM.SBJ=PLCT-PST-3ANIM.PL.SBJ-flew-**PL**
          'They flew (one after another).'

The use of reanalyzed adjective, action nominal, or aspect suffixes as plural or pluractional markers in Ket finite verb forms is sporadic and unpredictable. As already mentioned, such reanalyzed affixes differ most strikingly from canonical agreement morphology inherited from Proto-Yeniseian in occasionally expressing plurality of an inanimate class subject.

4 Semantics and discourse

The semantics of number as category values have been described in the relevant sections above. There is no additional data currently available to discuss the pragmatics of number in Ket or its role in Ket discourse.

5 Summary and conclusions

The preceding sections have covered all facets of number marking in modern Ket, offering observations on the historical evolution of plural marking patterns to help distinguish ancient features from more recent innovations. The core distinction between singular and plural in Ket encompasses most nouns and pronouns and is typically expressed by augmenting the bare singular stem with a suffix containing one of two nasal consonants: /n/ or /ŋ/. There is no evidence that plurals were ever productively generated by other means. Although some irregularities in plural formation, including ablaut and odd morphophonemic stem alternations, trace back to the proto language, they appear to be morphophonological or otherwise secondary. The grammatical dichotomy between animate and inanimate class nouns influences the choice of plural suffix allomorph and also dictates the form of 3rd person anaphoric pronouns. Because Yeniseian personal pronouns are the etymological source of possessive markers and verb agreement affixes, these systems likewise formally distinguish between inanimate and animate class entities. Due to its intrinsic connection with number marking, the grammatical factor of animacy thus casts an influence over virtually all areas of the morphosyntax in one way or another. By contrast, the subdivision within animate class nouns between masculine and feminine manifests itself in agreement morphology associated with singular nouns and pronouns only and may represent a Ket and Yugh innovation not traceable back to Common Yeniseian.

Both genetic and areal factors in the linguistic history of Yeniseian peoples enhance the value of Ket data for a cross-linguistic study of number marking. Yeniseian languages are genealogically isolated from all other families of northern Asia, and some features of Ket number marking are not found in any of the neighboring languages. This includes, in particular, the pervasive formal contrast evident throughout Ket morphology between how animate class and inanimate class nouns express and reflect number – a trait not observed anywhere else among the indigenous languages of Siberia. Other features, such as the use of plural suffixes on nouns is found widely in other indigenous language families of northern and Inner Eurasia.

In addition, the language's genealogical isolation makes contact-induced changes a particularly interesting topic of investigation. Ket speakers were the last hunter-gatherer-fishers in landlocked northern Asia. They continued to move about seasonally in small kin-based groups long after all of their neighbors had shifted to reindeer herding or, in parts of south Siberia, to stockbreeding. For centuries, demographic pressure induced the Ket hunting bands to take brides from pastoral groups as marriage partners who presumably had to learn the language as young adults. The influence exerted on Ket linguistic forms by young L2 speaking mothers from Turkic, Ugric, Samoyedic, and Tungusic speaking groups could possibly be implicated in the unusually high number of cases where plural marking developed

via semantic reanalysis in areas of Ket morphology where it had originally been absent. This includes the sporadic creation of dual number forms of nouns, the rise of multi-site subject plural marking on the basis of a moribund change-of-state affix in finite verb forms, pluractional marking developed from a reanalyzed action nominal suffix, as well as plural marking in a small subset of adjectives – a lexical class originally entirely devoid of inflections of any kind within the strongly head marking Yeniseian family. The expression of pluractionality in Ket could conceivably have been influenced by Selkup, with its rich inventory of *Aktionsart* forms that relate to quantification of events. Analogous features need not have been present as such in the neighboring languages originally spoken by Ket spouses for these L2 speakers to have influenced the spread of new patterns via semantic reanalysis. Among languages in contact with Yeniseian, adjective agreement, for example, is present only in Russian, and these patterns seem to have been established in Ket and Yugh already before the arrival of Indo-European languages in central Siberia.

Finally, it need be stressed that although it is now possible to describe Ket number marking in great detail and to account for many intricate aspects of its diachronic evolution, some facets of its morphological development still remain unclear. This particularly concerns the motivations behind certain irregular noun plurals, as well as the reason why only certain adjectives and action nominal stems developed number marking while most did not. The origin of multi-site subject marking in some Ket verb types also remains unclear.

Abbreviations

ABL	ablative case
ADES	adessive case
ANIM	animate class
DAT	dative case
F	feminine
INAN	inanimate class
M	masculine
NMLZ	nominalizer
OBJ	object
PL	plural
PLTC	pluractional
POSS	possessive
PRES	present tense
PRON	pronoun
PST	past tense
SBJ	subject
SG	singular

References

Bibikova, Vera S. 1976. O kategorii chisla v sisteme imeni prilagatel'nykh v ketskom jazyke [Number marking in Ket adjectives]. In Andrej P. Dul'zon & Tel'mina I. Porotova (eds.), *Jazyki i toponimija Sibiri* [Languages and toponymy of Siberia] 6, 90–96. Tomsk: TGPI.

Castrén, M. A. 1858. *Versuch einer jenissei-ostjakischen und kottischen Sprachlehre*. St. Petersburg: Imperatorskaja Akademija Nauk.

Georg, Stefan. 2007. *A descriptive grammar of Ket, part 1: Introduction, phonology and morphology*. Kent: Global Oriental.

Helimski, Eugene. 2016. S-singulatives in Ket. *Journal of language relationship* 14(3). 157–163.

Kotorova, Elizaveta & Andrey Nefedov. (eds.). 2015. *Comprehensive dictionary of Ket*. (2 vol.) Munich: Lincom Europa.

Nefedov, Andrey & Edward Vajda. 2015. Grammatical sketch of Ket. In Elizaveta Kotorova & Andrey Nefedov (eds.), *Comprehensive dictionary of Ket*. Munich: Lincom Europa.

Porotova, Tel'mina I. 1990. *Kategorija mnozhestvennosti v enisejskikh jazykakh* [The category of plural in Yeniseian languages]. Tomsk: Tomsk State University.

Porotova, Tel'mina I. 2002. *Slovar' govornykh form ketskikh sushchestvitel'nykh (s formami mnozhestvennogo chisla)* [Dictionary of dialectal forms of Ket nouns (with plural forms)]. Tomsk: TGPU.

Vajda, Edward. 2001. The role of position class in Ket verb morphophonology. *Word* 52(3). 369–436.

Vajda, Edward. 2004. *Ket* (Languages of the world/materials 204). Munich: Lincom Europa.

Vajda, Edward. 2013a. Metathesis and reanalysis in Ket. *Tomsk journal of linguistics and anthropology* 1(1). 14–26.

Vajda, Edward. 2013b. Vestigial possessive morphology in Na-Dene and Yeniseian. In Sharon Hargus, Edward Vajda & Danny Hieber (eds.), *Working papers in Athabaskan (Dene) languages 2012*, 79–91. Fairbanks: ANLC.

Vajda, Edward. 2017a. Ket incorporation. In Michael Fortescue, Marianne Mithun, & Nicholas Evans (eds.), *Handbook of polysynthesis*, 906–929. Oxford: Oxford University Press.

Vajda, Edward. 2017b. Patterns of innovation and retention in templatic polysynthesis. In Michael Fortescue, Marianne Mithun, & Nicholas Evans (eds.), *Handbook of polysynthesis*, 363–391. Oxford: Oxford University Press.

Vajda, Edward. 2019. Ket lexical typology. In Mikhail Kotkin (ed.), *Wortschätze und Sprachwelten* (Sprach- und Kulturkontakte in Europas Mitte: Studien zur Slawistik und Germanistik 10), 59–72. Berlin: Peter Lang.

Werner, Heinrich. 1997a. *Das Jugische*. Wiesbaden: Harrassowitz.

Werner, Heinrich. 1997b. *Die ketische Sprache*. Wiesbaden: Harrassowitz.

Werner, Heinrich. 2002. *Vergleichende Wörterbuch der Jenissej-Sprachen*. Wiesbaden: Harrassowitz.

Valentin Gusev and Beáta Wagner-Nagy

10 Number in Nganasan

Abstract: Nganasan (Samoyedic < Uralic) has a system of three numbers: singular, dual and plural, which are marked on all nominals for the object itself and, in the possessive forms, for the possessor. In the inflection, the three numbers are marked in all cases with the dual partly analytical. With paired objects, plural occurs more often than dual; in other cases when there are two objects, the dual is obligatory. Non-singular forms can have an associative reading. Numerals can also be in the plural, either referring to sets of objects or indicating an approximate quantity. Verbs have aspectual derivations which mark plurality of actions or arguments, such as iteratives, multiplicatives and so on. After numerals, nouns are used in the singular (except when the numeral itself is in the plural). Adjectives agree in number with their heads. Verbs agree in number with their subjects and, under certain circumstances, with the object.

1 Overview

Nganasan is a Uralic language belonging to the Samoyedic branch of the family and spoken on the Taimyr Peninsula in the north of Siberia. At present, most of the Nganasans have shifted to Russian. The change happened in the last quarter of the 20th century and was rather abrupt, therefore the language of the full speakers bears no visible traces of Russian influence on grammar. As far as we know, there are about a dozen full speakers left, all of them aged 70 or older, and some dozens of semi-speakers with various degrees of language proficiency.

This chapter is based on a corpus of texts recorded for the most part between the 1960s and the 2010s, and reflecting Nganasan as it was spoken by the last generations of its speakers. The corpus was annotated between 2015 and 2017 within the framework of the project *Corpus building and corpus-based studies of Nganasan*, supported by the DFG grant WA 3153/2-1. As part of this work, the Nganasan Spoken Language Corpus (NSLC) was compiled (Brykina et al. 2018[1]), now consisting of 21,723 utterances. In addition, we have used transcriptions of about 30,000 sentences that have not yet been glossed and are therefore not part of the published corpus. Our data consist mainly in folklore texts, but include other types of texts including narratives and conversations. The glossing in this study is occasionally simplified with respect to morphology, when it is irrelevant for the topic of this paper.

---

1 http://hdl.handle.net/11022/0000-0007-C6F2-8.

Nganasan, like its sister languages Enets, Nenets, and Selkup and neighbours such as Evenki or Dolgan, is an agglutinating language with a high degree of fusion and complex morphophonology. There are no noun classes or classifiers. Except for the morphophonological alternations, all nouns (except the numeral 'one' and the demonstratives *əmti* 'this' and *təti* 'that') and all verbs (except the verb 'be') are inflected regularly and uniformly, with no irregular or suppletive forms.

Nouns inflect for number, case, and possessivity. The category of the case includes three syntactic cases (Nominative, Genitive, and Accusative) and four local cases (Lative, Locative, Elative, and Prolative). In addition, there is a category of (pre)destinativity (meaning 'X designated for Y', see Daniel 2005, Leisiö 2014, or Nikolaeva 2015 and Khanina & Shluinsky 2014 about the related category in Nenets and Enets). Nouns also have predicative forms (meaning 'I am X', 'you are X' etc.). Adjectives agree with nouns in case and number.

Verbs inflect for tense and modality and agree in person and number with their subjects. They also have a special category, traditionally named "conjugation type", which shows the topicality of the object. When the object is known from the context, the verb takes the inflections of the so-called "objective conjugation", agreeing not only in person and number with its subject but also in number with its object.

Detailed descriptive grammars of Nganasan include Tereshchenko 1979, Katzschmann 2008 and Wagner-Nagy 2019; short descriptions with a particular emphasis on morphophonology can be found in Helimski 1994 and 1998.

2 Nominal and pronominal and verbal number

This section describes the formation and use of number forms in pronouns (2.2) and nouns (2.3) and provides some remarks about what can be related to the category of verbal number (2.4).

2.1 Generalities

Nganasan has three number values: singular, dual, and plural. All of them are obligatory, meaning that – except for some special cases that will be considered below – dual must always be used when speaking about two entities, and plural when speaking about entities that are more than two. The number is expressed in nouns and pronouns, and, by way of agreement, in adjectives, demonstratives and verbs. There is no obvious division of nouns into different classes relevant to the use of the number. In nouns and pronouns, the number values and their use are the same.

Attributes agree with their heads in number and, partially, in case. Verbs agree in person and number with their subjects, as illustrated by virtually any example

below. In some cases, they also agree in number with their object. See Section 3 on agreement in number.

There is no traditionally recognized category of "verbal number" distinct from agreement. However, among the many aspectual derivations, some denote the plurality of events, including Iterative, Frequentative, or Multiplicative. There are no productive singulative (semelfactive) derivations.

No dedicated verbal stems for singular or plural subjects or objects are known.

2.2 Pronominal number

As mentioned above, nouns and pronouns distinguish three numbers: singular, dual, and plural. Personal pronouns are homophonous in the nominative, genitive and accusative,[2] their number forms are suppletive, as shown in Table 1. Although the first and second person pronouns can be easily divided into two segments, one responsible for the person, the other for the number category, these segments are not used elsewhere in the language.[3]

The local cases of the personal pronouns are formed with the stem *na-* (otherwise a postposition 'on, above') with case and possessive affixes, e.g. *na-tə-nə na*-ABL-1SG '(away) from me', *na-nu-ntu na*-LOC-3SG 'near/with him', etc. The same stem *na* is used for local cases in the dual forms of nouns, see below.

Third-person pronominal reference is mostly performed by demonstrative pronouns, which have the same inflectional categories as the nouns. There is a kind of dedicated third-person pronominal stem, identical to the nominal stem meaning 'appearance' or 'shadow' *siʔ-*. In the pronominal function it is accompanied by the corresponding possessive marker, i.e. *sɨ-tɨ* shape-3SG 'he/she/it'; *sɨ-tɨ* shape-3DU 'they.DU'; *sɨ-tɨŋ* shape-3PL 'they.PL'.[4]

The demonstrative pronouns *əmtɨ* 'this' and *təti* 'that' have suppletive stems: *əmə-* and *təndə-*, *tətə-*, *təə-*, respectively; but their distribution is independent of number. In terms of inflection, other kinds of pronouns do not differ from nouns.

Tab. 1: Personal pronouns (NOM-GEN-ACC forms).

	sg	du	pl
1	*mənə*	*mii*	*mɨŋ*
2	*tənə*	*tii*	*tɨŋ*

2 Nganasan is typologically rare in distinguishing core cases in nouns, but not in personal pronouns.

3 Except that *-m* is one of the verbal agreement suffixes for 1SG. Historically, this is hardly a coincidence, but synchronically, one has to consider this as pure homophony.

4 These forms are rarely used and await further research.

2.3 Nominal number

2.3.1 Noun inflection

2.3.2.1 Non-possessive inflection

An unaffixed nominal stem is always grammatically singular; we are unaware of contexts where a bare stem would trigger dual or plural agreement on the verb.

The dual and plural forms of all nominals are formed in a regular way (taking into account the very complex but regular morphophonology), see Table 2. The numeral *ŋuʔəiʔ* 'one' has irregular stem variants *ŋuʔəð-* and *ŋuʔəj-*, but their distribution is independent of number.

The endings in brackets have now virtually disappeared, but do very rarely occur in an archaizing style.

It can be seen that in the Dual, only the three syntactic cases have synthetic forms; the local cases, on the contrary, are built with the case forms of the postposition *na* (the same stem that is used for personal pronouns, see above).

As concerns the other cases, it has been argued (cf. Däbritz 2017 among others) that in the plural, the case forms can be traced back to an "ideal" system with separate suffixes for the case and number, where the number is marked with *j* in oblique cases (although the plural accusative has no case marking, see Table 3). According to this view, the segments **ntə* and **kə* are the so-called "co-affixes"

Tab. 2: Non-possessive inflection of the noun *d'intə* 'bow'.

	sg	du	pl
NOM	*d'intə*	*d'intə-gəj*	*d'ində-ʔ*
GEN	*d'ində-[ŋ]*	*d'intə-gi*	*d'indi-ʔ*
ACC	*d'ində-[m]*	*d'intə-gi*	*d'ində-j*
LAT	*d'ində-tə*	*d'intə-gi na*	*d'ində-tiʔ*
LOC	*d'ində-tənu*	*d'intə-gi nanu*	*d'ində-tini*
EL	*d'ində-gətə*	*d'intə-gi natə*	*d'ində-gitə*
PROL	*d'intə-mənu*	*d'intə-gi naməənu*	*d'indiʔ-mənu / d'ində-mənu*

Tab. 3: Proto-North-Samoyedic inflection of nouns (SG and PL).

	sg	pl
NOM	**d'intə*	**d'intə-t*
GEN	**d'intə-n*	**d'intə-j-t*
ACC	**d'intə-m*	**d'intə-j*
LAT	**d'intə-ntə-ŋ*	**d'intə-ntə-j-ŋ*
LOC	**d'intə-ntə-nå*	**d'intə-ntə-j-nå*
EL	**d'intə-kə-tə*	**d'intə-kə-j-tə*

which are now parts of the local case suffixes but probably go back to some locational noun markers (cf. Mikola 2004: 98 ff.). The prolative case is a later development, so it shows a different structure. In the plural, the prolative suffix *-mənu*, which is of postpositional origin, is attached to the plural genitive stem.

The formation of Dual should have been similar to what it looks like now, though some details are not fully clear.

Cardinal numerals can inflect, too, including for plural (cf. Section 2.3.2.6), which is formed in the same way as for nouns.

2.3.2.2 Possessive inflection

Nominals also have possessive declension. The possessive indicates the person and number of the possessor and also the number of the possessed noun (see Table 4, the possessive declension of the noun *d'intə* 'bow').

As Table 4 shows, if only one entity is possessed, then the possessive series shown in column 1 (singular possessee) is used. For non-singular (dual or plural) possessees, another series (columns 2 and 3) is used. If two entities are possessed, the possessive suffix is preceded by the marker *-kəj/-gəj* expressing the dual possessee. In the plural paradigm, the same possessive series is used without the number marking for the possessee (etymologically, it contains the plural marker *-j-*, now fused with the possessive suffix). In Table 1, the difference between the nominal stem used in the singular and the dual versus the stem used in the plural results from morphophonological processes and is not per se a means to express number.

Possessive markers are cumulative with the case (in the nominative, genitive and accusative) or are placed after the case marker (in the local cases), see Table 5, where third person possessor forms are shown as an example.

Tab. 4: Possessive inflection of nouns in the nominative.

Possessor person and number	Number of the possessee		
	singular	dual	plural
1sg	d'intə-mə	d'intə-gəj-n'ə	d'indi-n'ə
2sg	d'intə-rə	d'intə-gəj-t'ə	d'indi-t'ə
3sg	d'intə-ðu	d'intə-gəj-t'ü	d'indi-t'ü
1du	d'intə-mi	d'intə-gəj-n'i	d'indi-n'i
2du	d'intə-ri	d'intə-gəj-t'i	d'indi-t'i
3du	d'intə-ði	d'intə-gəj-t'i	d'indi-t'i
1pl	d'intə-muʔ	d'intə-gəj-n'üʔ	d'indi-n'üʔ
2pl	d'intə-ruʔ	d'intə-gəj-t'uʔ	d'indi-t'üʔ
3pl	d'ində-ðuŋ	d'intə-gəj-t'uŋ	d'indi-t'uŋ

Tab. 5: Inflection of the possessed noun (3sg possessor).

	singular	dual	plural
nom	d'intə-ðu	d'intə-gəj-t'ü	d'indi-t'ü
gen	d'ində-tu	d'intə-gəi-tü	d'indi-tu
acc	d'ində-mtu	d'intə-gəj-t'ü	d'indi-t'ü
lat	d'ində-tə-tu	d'intə-gəi-tü na	d'ində-ti-tü
loc	d'ində-tənu-ntu	d'intə-gəi-tü nanu	d'ində-ntini-ntü
abl	d'intə-gətə-tu	d'intə-gəi-tü natə	d'intə-giti-tü
prol	d'intə-mənu-ntu	d'intə-gəi-tü naməənu	d'indi-mənu-ntu

2.3.2 Special uses of the nominal number

In this section, we will consider some uses of number forms specific to particular groups of nominals, such as body parts (2.3.2.1), mass nouns (2.3.2.2), and collective entities (2.3.2.3).[5] In 2.3.2.4, we will describe the so-called dyadic suffix, used mostly with kinship terms. Section 2.3.2.5 discusses the associative reading of non-singular forms. In 2.3.2.6 we consider the plural forms of numerals, and in 2.3.2.7 we discuss plural forms of verbal nouns.

2.3.2.1 Body parts

Nouns denoting paired body parts[6] are most often used in the plural (1–2), though dual is also possible (3–4).

(1) *S'elujkia-rə   təs'iəðə tə   bənü-Ɂə,   **s'ejmi-t'i**   kən'əðu-sa*
     male.reindeer-2sg PTCL   PTCL lie.down-PF eye-ACC.PL.3SG close-INF
     d'igə   n'i.
     mountain.GEN   on
     'The reindeer lay down on the mountain and closed its eyes'
     (TKF_990816_Lemming_flkd 38)

(2) *Ou,   ou,   n'i-kəə-ŋ.   **D'üða-t'ə**,   ŋuə-t'ə*
     INTERJ INTERJ NEG-ADMON-2SG.S arm-ACC.PL.2SG leg-ACC.PL.2SG
     t'ambi-l'i-ŋ.
     break-DUB-2SG.S
     'Oh, don't do [it]. You may break your arms, your legs.'
     (ChZS_080212_Djajku_flkd 22)

---

5 Units of measure, such as time units (*d'alɨ* 'day', *hüə* 'year') or distance units (*kərɨgəl'i* 'distance passed between two halts while wandering') do not exhibit any particularities.
6 Nganasan has one word for 'arm' and 'hand', and for 'leg' and 'foot'; these roots are glossed as 'arm' and 'leg', respectively, and translated according to the meaning of the example.

(3) *Tahari̭aa* **d'ütü-gəi-t'ə** *məkiʔa kəmə-ʔki-ʔə-m.*
PTCL   arm-DU-ACC.2SG behind hold-RES-PF-1SG.S
'I'll hold your hands behind.'
(JDH_00_TwoDolganBrothers_flk.520)

(4) *D'orə-ma-ntu* *kuntə* **s'ejmi-gəi-t'ü** *hon'ü-bi̭aði-təʔ.*
cry-NMLZ-GEN.3SG while eye-DU-PL.3SG swell-INFER-3PL.R
'While he was crying, his eyes swelled up.'
(JDH_00_TwoDolganBrothers_flk.338)

In the singular, these nouns may refer to singular organs, often accompanied by attributes such as 'one' (as in 5), 'which one', or 'left'/'right', etc.

(5) *Ŋuʔəiʔ nɨ* *ŋomtü-tü,* **s'ejmi-ðɨ** *ŋuə-l'aa.* **D'ütü-ðü** *ŋuə-l'aa,*
one woman be.sitting-PRS eye-3SG one-only arm-3SG one-only
ŋoj-t'ü *ŋuə-l'aa. Təti-rə barus'i kobtṷ̄ə.*
leg-3SG one-only that-2SG evil.spirit girl
'A woman is sitting, she has only one eye. One arm, one leg. This is an evil spirit's girl'
(MVL_080304_NjomuKamleguNy_flks.267)

(6) *Tə,* **d'üðü-mtü** *d'il'i-ʔə, sɨti təbtə nənsu-ʔi-ndɨ.*
PTCL arm-ACC.3SG raise-PF they.DU too stop-PF-2DU.R
'He raised his hand, and they two stopped, too.'
(ChND_080722_TwoFriends_flk 134)

Another use of the singular with body parts is to refer to the corresponding ability – e.g. to vision:

(7) **S'ejmi-ðɨ** *homi-tɨ.*
eye-3SG be.sharp-PRS
'He has a sharp eye.'
(CHKD_72_ManyTents_flk 180)

(8) *Tahari̭aa kərutətu* **d'üðü-tənɨ** *n'i-sɨðə-riʔ nəkərəbti-ʔ.*
PTCL   common arm-LOC NEG-FUT-2PL.O stop-CNG
'You cannot stop him with your bare hands.'
(JDH_060901_DolganChief_flkd 171)

As concerns the non-natural pairs, the use of the dual is, as far as we know, obligatory.

The fact that paired objects are mostly used in the plural and not in the dual suggests that Nganasan is another example of languages that distinguish between the dual proper (which expresses an accidental duality) and the ambal (used only for natural pairs, see Plank 1989: 308 ff): the latter seems to be less grammaticalized (non-obligatory) than the former.

2.3.2.2 Mass nouns

With mass nouns, the plural often refers to several individual instances of the substance in question, as in (9). Here the plural form of *biʔ* 'water' means portions of water that one person has to bring regularly. In (10), it refers to the water that a lot of people, who have just made tents, are making in each tent simultaneously.

(9) *Labsəkəə-ðuŋ tuu   tʼin-tə-ti,*   ***bið-əj***   *tojbutə-suðə bənsə,*
    youngest-3PL fire.ACC light-IMPF-PRS water-ACC.PL bring-FUT   all
    kobtũa-tu nʼi-sɨðə   maagəlʼitʼə melisi-ʔ.
    girl-3SG   NEG-FUT nothing   do-CNG
    'The youngest brother is making fire, bringing water, the girl will do
    nothing.'
    (JMD_080219_MyLife_nar 14)

(10) *Mað-u-tʼüŋ*   *mɨi-ʔə-ʔ,*   *hũaa-ði-tʼüŋ*   *hoidʼi-ʔə-ʔ,*
    tent-EP-ACC.PL.3PL make-PF-3PL.S wood-DEST-ACC.PL.3PL chop-PF-3PL.S
    bi-ti-tʼiŋ   *meliði-ʔə-ʔ.*
    water-DEST-ACC.PL.3PL make-PF-3PL.S
    'They made tents, chopped wood, made water [from ice].'
    (MVL_97_Tangka_flkd 6)

In (11), the same form (*biðəʔ*, plural of *biʔ*) refers to rivers and lakes. The expression "waters moving" is the usual way to denote the breaking up of the ice in the spring (that is, in the first days of June).

(11) *Sʼüəbtĩaiʔ ŋanuə noruə isʼüə,*   ***bið-iʔ***   *büü-tə-büʔə.*
    really   PTCL spring be-PST water-GEN.PL go-IMPF-VN
    'It was in the spring when the ice broke up on the rivers.' (lit. "at the waters'
    being about to move").
    (PTK_93_Djemnime_nar 20)

The same word *biʔ* may also mean 'alcohol', and then its plural form refers to bottles or cups, as in (12), or to different sorts of alcohol, as in (13).[7] (See also example (26) for the associative reading of the same form.)

(12) ***Biði-tʼi***   *təniʔariaiʔ tutu-ti-ʔ.*
    water-PL.3PL just.so   stay-PRS-3PL.S
    'Cups (of vodka) are [on the table].'
    (TKF_061105_FoxFosterling_flk 298)

---

7 The interpretation of examples (12) and (13) is based on the translations given by the informants.

(13) *Təti-rə   kuʔ ŋanuə t'üünə-j   **biðə-j**   ŋanuə təndə   mou*
    that-2SG very PTCL different-ACC.PL water-ACC.PL very that.GEN place
    n'i-məəni n'enat'aʔa kəbüra-ʔa   tatə-ra-ʔa,   təndə   n'i-məəni
    on-PROL huge   carpet-AUGM spread-PASS-PF that.GEN on-PROL
    ho-la-ʔi-ndəʔ.
    put-PASS-PF-3PL.R
    'So many different sorts of alcohol were put on the ground, then they spread
    a big carpet and laid everything onto it.'
    (KES_031115_Paris_nar.245)

Here are some examples of other mass nouns. (14) may be interpreted as an example
of the abundance plural, and (15) and (16) as having a distributive reading.

(14) *Ŋuʔəiʔ numajkaʔa taharĩaa ŋanuəməni kurəgüiʔ ton'd'ajt'i-sɨ **horba-t'ü***
    one   boy   PTCL   PTCL   PTCL   lie-INF   foam-PL.3SG
    ŋəðu-tɨ.
    be.seen-INF
    'A boy is lying, foam is seen [on his lips]' (lit. "his foam[s] are seen."
    (TKF_990812_EvilSpirit_flkd.274)

(15) *Kət'i kolɨ-ə   **d'irə-ʔ,** tərəd'i-ʔ i-kə-d'üə-ʔ.*
    PTCL fish-ADJ fat-PL   such-PL be-ITER-PST-3PL.S
    [At that time we made lamps with fat, you know. With polar foxes' fat] 'And
    [with] fish fat, there were such [lamps]'
    (KES-ChND_080725_Childhood_conv.160)

(16) *Lakunaʔa   ŋəndiãiʔ, mabtad'a **kamu-t'ü**   sohə-nar-ŋa-ntəʔ.*
    having.tuberculosis probably why   blood-PL.3SG fall-FRQ-INTERR-3PL.R
    'He must have tuberculosis, that's because he's spitting blood' (lit. "his
    blood[s] are falling").
    (ChND_080722_TwoFriends_flk.131)

Example (17) contains names of several different metals (the first two are unknown
to modern speakers and are translated tentatively) in the plural, referring to the
parts of a bow made from them.

(17) *Əmɨʔa tolami n'aagəiʔ sol'üʔ-tüə   d'aðüd'ə basa i-huãðu,*
    so   mixed well   shine-PRT.PRS forged iron be-INFER
    ***horaŋgana-ʔ, tulegu-ʔ, s'eribiri-ʔ, basaðaŋə-ʔ, n'orumu-ʔ** – tərəd'i-ndini*
    cast.iron-PL steel-PL silver-PL brass-PL   copper-PL such-LOC.PL
    bənsə kəlbiʔ-məə.
    all   make.stripes-PRT.PASS
    '[The bow] has many kinds of shining forged iron [= metal on it]: cast iron,
    steel, silver, brass, copper – it was all made with strips of these'
    (SM 1226)

Dual is possible with mass nouns as well, cf. example (18):

(18) *Təti mənü-kəi-t'ə tə   maagəl'it'əgətə malamsəə-gəj **basa-gəj** təʔ.*
     this fist-DU-2SG PTCL very   different-DU iron-DU PTCL
     'Your two fists are made from different [kinds of] iron'
     (132_KTD_87_IronHead_flks 1199)

2.3.2.3 Collective entities

With some nouns, their plural forms are a standard way to refer to an entity that consists of several elements denoted by these nouns. Such use is standard with three words: *maʔ* 'tent', *koruʔ* '[wooden] house', and *muŋku* 'tree'. Their plurals can mean 'nomad camp' (*maðəʔ*), 'village' (*koruðəʔ*) and 'forest' (*muŋkuʔ*), respectively; there are no other words for the camp, village, or forest. At the same time, the plain plural uses meaning (several) 'tents', 'houses' and 'trees' are also possible. The use is not limited to these three nouns. For instance, the plural form of the old Russian loanword *l'ekɨr* 'doctor' can mean 'hospital' (and this is the only way to name it). Compare the following two sentences. In (19) the plural form of *l'ekɨr* means 'hospital', while in (20) it means just 'doctors'.

(19) *Təti tənəd'ad'ə i-s'üə, bəŋkə-mə tamnu **l'ekɨr-uʔ**   kad'anɨ i-s'uə.*
     that summer   be-PST balok-1SG there doctor-GEN.PL near   be-PST
     'It was summer, my *balok* [mobile dwelling put on a sledge] was there near the hospital.'
     (JMD_080219_MyLife_nar 174)

(20) *Təniʔaj-t'ü-tə-nuʔ   tahari͡aa maaŋuna ŋanuəmənɨ, tə-tə*
     be.SO-NMLZ-LAT-2PL PTCL   what   really   PTCL
     *skorɨj-ŋalə-ʔ   maa-ŋalə-ʔ **l'ekɨr-əʔ** n'i-ŋɨ-ʔ   tuu-ʔ*
     ambulance-PTCL-PL what-PTCL-PL doctor-PL NEG-INTERR-3PL.S come-CNG
     [An old woman slipped and fell.] 'While we were [bustling around her], the ambulance and doctors arrived.'[8]
     (KES_031115_Paris_nar 218)

2.3.2.4 Dyadic suffix

Dyadic, also called "connective-reciprocal" (the term first proposed by Hajdú 1975: 71–114), refers to the suffix, present in some Finno-Ugric and Samoyedic languages, which is added mostly to kinship terms and expresses mutual relationship. The

---

8 Negative interrogative sentences in North Samoyedic are often used for affirmative sentences: see Wagner-Nagy 2019: 424–425 and Gusev 2020.

most frequent word with this suffix in Nganasan folklore (where its forms are *-d'ə*[9] or *-səd'ə*) is *n'ini-d'ə* older.brother-DYA 'two brothers', literally "older brother and younger brother(s)' as in (21) and (22). This suffix can also be used with *ŋahu* 'older sister' (*ŋahud'ə* 'sisters' or 'older sister with brother(s)'), *n'emi* 'mother' (*n'emid'ə* 'mother with her child(ren)'), *d'esi* 'father' (*d'esid'ə* 'father with his child(ren)'), *ni* 'woman, wife' (*nisəd'ə* 'spouses, husband and wife'), *n'a* 'companion, partner' (*n'asə*[10] 'partners, friends, etc.'). Note that the dyadic stems are always based on a root which refers to the older member of the group: father or mother, older brother, or older sister. For 'spouses', only the word *nisəd'ə* is used; the parallel form based on the root *kuəd'umu* 'man' is not attested and seems not to exist.

(21) *S'iti* **n'ini-d'ə,**   *as'a-gəj,*   *as'a-gəj*   *tə-s'iə-gəj,*   *d'aŋur-uʔ*
   two brother-DYA Dolgan-DU Dolgan-DU that-PTCL-DU tundra-GEN.PL
   n'ii-məni   *d'otür-ü-gəj.*
   direction-PROL walk-PRS-3DU.S
   'Two brothers, Dolgans, these Dolgans, are wandering in the tundra.'
   (JDH_00_Fosterling_flkd 2)

(22) *Nagür* **n'ini-d'ə**   *i-śüə.*
   three   elder.brother-DYA be-PST.3SG
   'There were three brothers.'
   (ChNS_080214_Sjunezi_flks.002)

When these terms are preceded by a numeral, for example, 'two' as in (21) or 'three' as in (22), they are not marked for number, following the general rule. Without a numeral, a number marker corresponding to the "resulting" number of persons is added to the stem with a dyadic suffix, e.g. *n'ini-d'ə-gəj* older.brother-DYA-DU 'two brothers', *n'emi-d'ə-gəj* mother-DYA-DU 'mother and her son/daughter', *n'emi-d'ə-ʔ* mother-DYA-PL 'mother and her children (two or more)', etc. Therefore, the dyadic category may be thought of not as a primary number category, but as a conceptual category of kinship groups, which only implies non-singularity but does not include it as part of its grammatical meaning.

   For agreement, see Section 3.

2.3.2.5 Associative plural

Dual and plural forms in Nganasan may have an associative reading. This reading is very frequent with the nouns *n'emi* 'mother' and *d'esi* 'father', which in the dual

---

9 A homonymic or etymologically the same suffix with roots expressing qualities means 'of the same (size, height etc.)': *təŋkə-d'ə* 'equally strong' (*təŋkə* 'strength'), *hirə-d'ə* 'equal; of the same age' (*hirə* 'level'), and so on.
10 This form is not found in the texts, but is attested in Helimski's unpublished lexical data on Nganasan (Helimski ms.).

and plural mean 'parents'. Dual and plural can be used interchangeably, and the two words may be used together or separately; cf. two examples from the same speaker:

(23) *Mənə **n'emi-n'ə** **d'ed'i-n'ə** ŋünĩa-ʔ i-s'üə-ʔ, manu hün's'ərəənu.*
I mother-PL.1SG father-PL.1SG rich-PL be-PST-3PL.S in.the.old.days
'My parents were rich, in the old days.'
(JSM_090809_ParentsAndUncles_nar 1)

(24) *Taharĩaa kobtũə ŋomtü-tü **n'emi-gəi-ti** nanu, n'emi-nti*
PTCL girl be.sitting-PRS mother-DU-GEN.3SG with mother-GEN.3SG
kad'anu.
near
'The girl is sitting with her parents, near her mother.'
(JSM_060901_Relationship1_nar 2)

For pragmatic reasons, the regular (additive) plural reading of these two nouns is rare, but not impossible. For (25), we know from the translation given by the narrator herself that it means not 'parents', but 'fathers' (also, tooth carving typically was men's work, not women's).

(25) *Təgətə ŋonəə **d'ed'i-t'iŋ** mej-kə-tiʔ əligaku-j*
then also father-PL.3PL make-ITER-PRS-3PL.S small-ACC.PL
kənda-ʔku-j, kunsɨbĩa-ʔku-ʔ i-gə-tu-ʔ, in's'üðə-ʔ
sledge-DIM-ACC.PL load.sledge-DIM-PL be-ITER-PRS-3PL.S man.sledge-PL
kənda-ʔku-ʔ.
be-ITER-PRS-3PL.S
[People used to gather reindeer's teeth, as toys for their children.] 'Then their fathers made small sledges, they were sledges for loads, sledges for men.'
(ChND_041213_Reminiscence_nar 141)

The associative plural is possible with other nouns as well:

(26) *...takəə **d'intə-gəl'it'i-n'ə** koi-ʔə-m kənda-tə-nə.*
that.far bow-EMPH-ACC.PL.1SG leave-PF-1SG sledge-LAT-1SG
[A man saw a bear and thinks: "What will I shoot at him with,] 'I've left my bow and arrows on the sledge.'"
(JDH_00_TwoDolganBrothers_flk 36)

Since a man normally has only one bow, a non-associative reading of this example is improbable.

Cf. also the following example. Here, bringing water is regarded as one of the domestic duties, and, correspondingly, the plural form of 'water' refers to the various household tasks.

(27) *Tə   ŋuəli   **bɨði-n'ə**   melisi-ku-ðəum   baðətə-nə.*
PTCL   of.course water-ACC.PL.1SG do-IMP-1SG.S.EXCL self-1SG
'Of course, I'll bring myself water and do other household chores.'
(JDH_99_ThreeTents_flkd 129)

Finally, an associative reading of plural or dual is also possible with proper names. In (28), the Dual form of the proper name refers to the man and his fishing companion– in this case, not a member of his family, though this would also be possible.

(28) *Əmtɨ, L'on'a-rə mɨəd'ɨ-sɨ   d'öðür-kəə?   – Təti-rə   **Nagorn'uka-ʔa-gɨ***
this   L.-2SG   go.by.foot-INF go-ITER.INTERR   that-2SG N.-AUGM-GEN.DU
nanu kona-ʔa.
with go-PF
'This L., your (son), did he go [fishing] by foot? – [No], he went with N. and his companion [on snowmobile].'
(KES-SEN_031115_Dialog_conv 02:25)

2.3.2.6 Plural form of the numerals
The numerals themselves can be in the plural; in these cases, the nouns they are related to are in the plural, too. Such constructions can have two possible readings. First, they can denote several groups of objects, each of which consists of the specified number of entities, that is 'by two', 'by three', etc.

(29) *Bən'd'ikaa-j kotə-ʔsuði-t'i,   **ŋuʔəðə-j**   ŋanaʔsanə-j,   **sɨði-j***
all-ACC.PL   kill-FUT-3SG.OP one-ACC.PL person-ACC.PL two-ACC.PL
ŋanaʔsanə-j   kotə-ʔsuði-t'i.
person-ACC.PL kill-FUT-3SG.OP
'He'll kill all of them, he'll kill them by one, by two.'
(MVL_090809_NarrowFacedGirl_flks.579)

(30) *Əmi-ʔ   d'alɨ-tini,   **kanə-j**   **biiðə-j**   hüə-j*
this-GEN.PL day-LOC.PL how.many-ACC.PL ten-ACC.PL year-ACC.PL
bəu-tu-tə-nuʔ,   batuðü-t'ü   mütəmi-ʔə.
pass-NMLZ-LAT-1PL, messenger-ACC.PL.3SG send-PF
'In these days, after we lived for several decades, he sent messengers.'
(ChND_061025_Haljmira_flks.43)

Second, phrases with a numeral in the plural can be translated as "approximately".[11] Such use is attested with expressions for hundreds (one hundred, two hundred, three hundred) and, once, for tens (which is more natural for the approximate meaning than exact cardinalities such as seven or thirty-eight). The construction often occurs in folklore texts when describing a huge quantity of tents or reindeer. It may be relevant that in most cases it is used to describe huge quantities, unimaginable in reality (in folklore texts, as in 31), or just denote a big sum of money, as in (32):

(31) **Nakü͡rə-ʔ d'irə-ʔ**   *maðə-ʔ. Təgətə taha͡rĩaa maðu-ʔ*   *əmənikani*
three-PL hundred-PL tent-PL then   PTCL   tent-GEN.PL aside
ŋuʔəiʔ maʔ nən'd'i-ti.
one   tent stay-PRS
[There are about] three hundred tents. Aside from the tents, there is one
[more] tent.'
(JDH_99_ThreeTents_flkd.3–4)

(32) *Mənə taha͡rĩabə n'üə-nə*   *basa*   *i-s'üə* *t'etə basa. Timĩ͡nia*
I   PTCL   child-GEN.1SG money be-PST four money now
taha͡rĩabə hiəməu, kanə,   **mətü͡ʔ d'irə-j**   *bad'a-j*
PTCL   INTERJ how.many six   hundred-ACC.PL money-ACC.PL
kəmü͡ðü-ndü-ʔ
receive-PRS-3PL.S
'My child allowance was 4 roubles. And now, people say, they get, how
much, six hundred roubles.'
(KES-ChND_080725_Childhood_conv.213–215)

2.3.2.7 Action nouns

Nganasan has two nominal forms that are regularly formed from every verb. They are traditionally called *perfective* verbal noun *-ʔmuə* and *imperfective* verbal noun *-mun*. (Their use is described in detail in Gusev 2012: 350–357.) The imperfective nominal form refers to the action as such, without mentioning its limits; among the typical uses of its case forms are 'while doing smth', 'unless smth occurs' or 'wish smth be done'; in our corpus, it is never used in the plural, except for its use to describe sounds, see below.

The perfective nominal form, on the other hand, denotes the entire action, including its endpoint and often its result. Here, the plural use is quite typical. Note that in many cases, the fact that there is a single event consisting of several phases explains the use of the plural.

---

11 Sometimes the informants say that these constructions mean just the same as their counterparts in singular.

(33) *Tahaȓi͡abə təndə   kobtu͡ā təndə   **hu͡aŋkəða-ʔmuə-d'əə-j**   bən'd'ə*
PTCL   that.GEN girl   that.GEN kidnap-NMLZ.PF-ANT-ACC.PL all.ACC
d'ebtaða-tu.
tell-PRS
'He tells everything about how he kidnapped this girl.'
(JSM_080217_FourBrothers_flkd 124)

(34) *Tə   timini͡a tahaȓi͡abə ŋanasanu-ʔ   **s'üd'a-ʔmui-ʔ**   t'ühə-nɨ*
PTCL now   PTCL   person-GEN.PL finish-NMLZ.PF-GEN.PL time-LOC
əmə-ʔ l'üəsa-ʔ   tahaȓi͡abə tuu-na-l'ɨ-ʔə-ʔ   əmd'ümə kəi d'a
this-PL Russian-PL PTCL   come-MULT-INCH-PF-3PL.S this.GEN side all
əbəjd'aʔi-ʔ   mou-gitə.
frightful-GEN.PL land-ABL.PL
'Now, when [our] people are almost over, Russians began to come here from
distant lands.'
(KES-ChND_080725_Childhood_conv 128)

(35) *Tə-tə əndɨ-t'iʔ,   **vistupaira-ʔmüə-d'əi-t'üʔ**   t'etu͡ā*
PTCL what's.its.name-PL.2PL perform.R-NMLZ.PF-ANT-PL.2PL very
n'aagəə-mənu višl'i.
good-ADV   proved.PL.R
'Your performance proved very good.'
(KES_031115_Paris_nar 14)

(In the latter example, it is known from the context that there had been only one
performance, but several performers had taken part in it; apparently, this explains
the plural form.)

Another deverbal noun, which is also very productive, has the suffix -(*b*)*sa*(*n*).
These nouns may refer to instruments or places; they are also often used for creating
nouns for objects that did not exist in the traditional life: cf. *s'eŋəbs'a* 'place for
passing a night' (*s'eŋə-* 'spend a night'), *ŋomtüs'a* 'chair' (*ŋomtü-* 'be sitting'), etc.
Such nouns can be easily used in the plural. Cf. also *səŋürs'an-əʔ* PL 'binoculars',
which may be an occasional formation from *səŋür-* 'look', but is a pluralia tantum
(despite the fact that Russian *binokl'* is singular).

Nouns denoting sounds, especially repeating ones, tend to be used in the plu-
ral. This is typical for such nouns as e.g. *d'el'i* and *saü* 'noise'. *Mirəimüʔ* 'noise of
footsteps' is only attested in the plural. *Buəmunəʔ* 'human voice, noise of conversa-
tion' is also used mostly in the plural. It is an imperfective nominal form of the verb
buə- 'speak', and this is the only known exception of what was said above about
the non-use of this form in the plural. It can be hypothesized (see Gusev 2017) that
such forms have given rise to the contemporary auditive evidential forms in the
modern Northern Samoyedic languages. The auditive evidential denotes events per-

ceived non-visually (most often through the auditory channel, but occasionally by
other senses as well).

2.4 Verbal number

Among the large number of aspectual derivations in Nganasan, there are some that
indicate the multiplicity of actions or participants, for instance the iterative, distribu-
tive and multiplicative (see more details in Gusev 2012, Wagner-Nagy 2019: 532–
537). While the iterative in (36) refers to different instances of an event with the
same participants, the distributive in (37) and the multiplicative in (38) also indicate
the plurality of subjects and/or objects; however, the distribution between them is
not yet fully clear.

(36) *Sili-r͡iai-ʔ* *maðu-ʔ* *d'a* **kontu-ru-gə-ta-ðə.** *ƏʔƏ,*
 who-LIM-GEN.PL tent-GEN.PL ALL take.away-PASS-ITER-PRS-3PL.R yes
 s'itəbɨr-n͡iagə-tɨ. *S'itəbɨr-n͡iagə-tɨ.* *ƏƏʔ.* *In's'üðü-tüə-ŋalə-ʔ*
 tell.tales-SUP-3SG tell.tales-SUP-3SG yes go.on.sledge-PTCP.PRS-EMPH-PL
 tuj-kə-tu-ʔ, *sɨtɨ huud'i-tɨə-ʔ.* *Vələt'anka d'a*
 come-ITER-PRS-3PL.S he go.for-PTCP.PRS-PL Volochanka ALL
 büü-bü-tü, *Vələt'ankə-ndə* **kəmə-ru-gə-ta-ðə** *ŋanaʔsa-ndiʔ.*
 go.away-COND-3PL Volochanka-LAT catch-PASS-PRS-3SG.r person-LAT.PL
 [Your father was a famous story-teller and singer. Yes, he was. I know him.
 Sure. He has just left us.] 'People used to take him to their tents. Yes, to hear
 his tales. To hear his tales. Yes. They even came on their sledges to take him
 with them. If he went to Volochanka [the nearby village], people used to
 catch him in Volochanka.'

(37) **Ŋəmə-bt-alu-d'üə-mə** *kuniʔa təniʔa, buə-tu.*
 eat-CAUS-DISTR-PST-1SG.O how so say-PRS
 ['Her husband came back: "Have you had a guest?" – "Yes, the guest came".]
 'I served her different things," she said.'
 (ChND_041212_Girl_flkd.176)

(38) *Tə, taluə* *d'alɨ-gətə n'endu hunsəə-j bitið-əj*
 well yesterday-ADJZ day-ABL precisely other-ACC.PL arrow-ACC.PL
 nəmtə-ʔnar-u-m
 discover-MULT-PRS-1SG.S
 'Oh, since yesterday I have been discovering foreign arrows.'
 (JDH_00_FallenEarth_flkd 48)

There are no special verbal stems for singular or plural subjects or objects.

3 Agreement and the syntax of number

Attributes (incl. numerals) usually precede the head noun. After numerals, nouns typically are in the singular (39). Occasionally, however, after the numeral 'two', the noun can take the dual (40) or plural marker. After the numerals greater than two, the plural marker can also be occassionally used, cf. (41), where singular and plural are used interchangeably. The distribution of the number forms after numerals is not clear.

(39) *Tahar͡iaa təti* **s'iti numajkaʔa,** *ŋuəlɨ*   *manakü bəðu͡a-tə-tuə-gəj*
     PTCL   that two boy   of.course just   grow-IMPF-PTCP.PRS-2DU
     numajkaʔa-gəj taat'üguj-hu͡aŋhu-gəj   *hiin'd'a.*
     boy-DU   herd.reindeer-RENARR-3DU.S in.the.night
     'Now, these two boys, two just grown-up boys, were herding reindeers in the night.'
     (ChNS_080214_Hotarye_flkd.015)

(40) *Tə,*   *təni͡ʔia* **s'iti əm-kəj numajkaʔa-gəj** *s'iti i-s'a*   *n'a-ti*
     PTCL so   two this-DU boy-DU   two be-INF one.another-GEN.2PL
     kad'a, n'a-ti   *n'iri*   *nənsu-ʔi-ndi.*
     near   one.another-GEN.2PL abreast stand-PF-3DU.R
     'So, these two boys, the two of them, stand up side by side.'
     (ChND_080722_TwoFriends_flk 275)

(41) *Tahar͡iaa* **s'ajbə kontə**   **taa**   *ŋanaʔsanu-n'ə koðuʔ-suðə-ʔ.* **S'ajbə**
     PTCL   seven sacrifice reindeer person-PL.1SG kill-FUT-3PL.S seven
     kontə   **taa.**   *Tahar͡iaa əmti* **s'ajbə baŋ,** *kunini tərəd'iʔ banə-ʔ*
     sacrifice reindeer PTCL   this seven dog where such-PL dog-PL
     təi-ŋu-ʔʔ   **S'ajbə banə-ʔ, s'ajbə kondə-ʔ**
     be.present-INTERR-3PL.S seven dog-PL seven sacrifice-PL
     təi-ŋu-ʔʔ   *...* **S'ajbə kontə,   s'ajbə taa,   s'ajbə baŋ.**
     be.present-INTERR-3PL.S   seven sacrifice seven reindeer seven dog
     'My men will sacrifice seven reindeer. Seven sacrificial reindeer. These seven dogs, where are these dogs? Are there seven dogs, seven sacrifices? ... Seven sacrifices, seven reindeer, seven dogs.'
     (MVL_090807_Bebtie_flk 313–324)

Remember that when the numerals themselves are in the plural, the nouns after them take the plural marker, too (Section 2.3.2.6).

Obviously, sets of two or more entities may be referred to simply with a dual or plural form, not necessarily with a numeral phrase. In the case of duals, the two options convey the same meaning. The question arises, what governs this choice?

It seems that at least one of the factors is the status of the information. If the objects in question are mentioned for the first time, they are used with the numeral;

when they are already known, dual or plural is enough. Cf. (42): the two sledges, when mentioned for the first time, are described as "two sledges". In the following sentence, they are referred to just as "sledges-DU". Then there are the young men sitting in them who are described as "two boys", and then, after a while, they, in their turn, also become simply "boys-DU". And as the young men already know their brides, they also refer to them simply with a dual demonstrative.

(42) *Təniʔia ŋomtü-l'aa-tu-ndə-ti   tahañaa* **s'iti in's'üðə-ʔa-gümü-rə**
     so   be.sitting-only-NMLZ-LAT-2SG PTCL   two sledge-AUGM-EMPH-2SG
     n'i-ŋɨ   tujhûa-ʔ.   **In's'üðə-gəj** *tuu-ʔə-gəj,   nənsu-ʔi-ndɨ.*
     NEG-INTERR approach-CNG sledge-DU   come-PF-3DU.S stop-PF-3DU.R
     Maaŋuna   tahañaa **s'iti numajkaʔa.** ... *Tə*   **numajkaʔa-gəj**
     what's.this PTCL   two boy   PTCL boy-DU
     munu-ntu-gəj: "Ou, **taaniə-gəj** *tuu-ʔə-gəj   təʔ,* **n'akiði-tiə-gəi-n'i***".
     say-PRS-3DU.S oh   that-DU   come-PF-3DU.S PTCL take-PTCP.PRS-DU-1DU
     'Two sledges are approaching [the camp] while they were sitting. The sledges came and stopped. What's this – two young men [are in the sledges] ... The young men say: "Ah, our brides ["those whom we are taking"] have arrived."'
     (ChND_080729_Mosquitos_flkd 144–146).

However, there are other factors as well. In folklore texts there are often groups or characters (for instance, brothers) who are called by their clan name e.g. *S'iti S'ünəðiʔia* 'two Syunazya', or just "three brothers" etc., who form a narrative unit; the same can be true for reindeer, like "three white reindeer". Such nominations are persistent throughout the whole text.

As for the attributes, there is an agreement in number and partial agreement in case between the attribute and the head noun. The head governs the attribute, thus if the head is in the dual or plural, the attributive takes the same number marking.

As far as the case marking is concerned, only grammatical cases (nominative, accusative, genitive) are marked on the attributive. In all other cases, the pre-modifier appears in the genitive case, as in (43). The modifiers do not agree in possessiveness.

(43) **Anikaʔi-ʔ d'ali-raa-tini**   *tətu-ru-gə-tu-ndəʔ*
     big-GEN.PL day-LIM-LOC.PL bring-PASS-ITER-CO-3PL.RFL
     bɨ-ti-ńiʔ   anikaʔa d'ali   i-hüʔ   tahañabiʔ.
     water-PL.DST-PL.1PLPOSS big.GEN day.GEN be-COND now
     'We were only given vodka on holidays[12], when there was a holiday.'
     (ChND-080719_Life.031)

---

[12] *Anikaʔa d'alɨ* is originally a traditional feast. In modern Nganasan this word is used with the meaning 'holiday'.

When nouns are used predicatively, they typically agree with the subject by taking person-number suffixes identical to those of the intransitive verbs, in which case they do not take number marking of their own. (In the third person, the dual and plural markers on nouns and intransitive verbs are identical.) In the present tense, the copula ('be') is omitted. The person-number suffixes may actually appear on the noun or on the copula alone, but, in careful speech, they appear on both.

(44) *Mənə les'n'iki-ʔ* **baarbə-m,** *les'n'iki-ʔ*
     I   forest.guard.R-GEN.PL chief-1SG.S forest.guard.R-GEN.PL
    brigad'ir-ə-m.
    group.leader.R-EP-1SG.S
    'I'm the chief of the forest guards, foreman of the forest guards.'
    (KTD_XX_MyLife_nar.022)

(45) **Ŋukəgəi-rbiʔa-miʔ i-s'üə-muʔ** *d'aŋur-u-ʔ*   *n'ii-ni*   *n'üə-ʔ.*
    many-AUGM-1PL   be-PST-1PL tundra-EP-GEN.PL on-LOC child-PL
    'There were many of us, children, in the tundra.'
    (KECh_080214_Childhood_nar.txt 4)

The predicative nominal construction does not distinguish between, e.g., 'we are such teams' and 'we are such a team'. Compare the following two examples: in (46), the predicate nouns ('persons' and 'workers') refer to two individual entities; in (47), to one collective entity.

(46) *Nəŋhə* **ŋanaʔsa-ʔku-mi** *i-t'ü-tə-ni,*   *n'intuu ŋojbəu-ʔs'a-ʔku-mi.*
    bad   person-DIM-1DU be-NMLZ-LAT-1DU NEG   work-NMLZ.ACT-DIM-1DU
    '... since we two are bad people, not workers.'
    (KES_080721_Disease_flkd.220)

(47) *N'aa-n'üʔ*   *əmə-ʔ Schmidt* **kolhosə-muʔ**   **i-s'üə-muʔ.**
    Nganasan-PL.1PL this-PL Schmidt collective.farm-1PL be-PST-1PL.S
    'Us Nganasans, we were [in the] collective farm [named after] Schmidt' (lit.
    "these Nganasans of ours, we were the Schmidt collective farm").
    (TND-PMM_061012_Dialog_conv.160)

Verbs agree in person and number with their subjects. They also have a special category, traditionally named "conjugation type", which shows the topicality of the object. When the object is known from the context (and most often omitted), the verb takes the person-number endings of the so-called "objective conjugation" and, in addition to the agreement in person and number with its subject, also agrees in number with its object.

(48) *Maa is'a* **kondüʔ-tə-tə-ŋa-t'ə?**
    what be-INF take.away-IPFV-FUT-INTERR-2SG.op
    'What do you take them away for?'
    (JSM_090809_ParentsAndUncles_nar.109)

(49) *N'aagəiʔ* **nogula-ʔa-gəi-t'ü.**
    well   approach-PF-DU-3SG.op
    'He came close to these two.'
    (TKF_031117_ThreeBrothers_flkd.048)

4 Semantics and discourse

In this short section, we consider some groups of nouns that tend to be used either in the singular or in the plural. There are no nouns that prefer the dual, except for the words for 'mother' and 'father' in the meaning of 'parents', and even with these two nouns in this meaning the plural is also possible; see Section 2.3.2.5.

4.1 Singularia tantum

This class is barely attested in Nganasan. It seems that virtually any word can take plural affixes. So far the word for 'mist' *kəku* is only attested in the singular, but its near-synonym *kəsɨ* 'haze, mist' can be used in the plural. *D'alɨ* 'day' is used only in the singular when it means 'weather'. Words denoting unique objects such as *kou* 'sun', *ŋuə* 'sky', *kit'əðəə* 'moon' occur only in singular, but *ŋuə* as 'deity' and *kit'əðəə* as 'month' (and also *kou* 'ear', the homonym of *kou* 'sun') have the plural forms.

4.2 Pluralia tantum

Some words are only or predominantly used in the plural. Strict pluralia tantum are *kəbi-a-ʔ* 'shaman's costume' (which consists of several parts), *kamar-əʔ* 'waist, loins' (cf. English *loins* or French *reins*) and a compound name *s'eimɨ kunsin-əʔ* 'eyeglasses' (the first part of it means 'eye', the second is not used on its own). The plural form is typical for such words as *kaðɨ-ʔ* 'scales', *n'erbɨ-ʔ* 'animal hair', *ŋabtə-ʔ* 'human hair', *kəəlɨ-ʔ* 'tears'.

    The plural form is also typical for the stem *d'aŋur* 'tundra' when it is used as a cardinal direction 'north-west', as opposed to *muŋku-ʔ* 'forest' ('south-east').[13]

---

13 On the Taimyr Peninsula, the border between the forest and the tundra goes from south-west to north-east, and this determines Nganasans' system of cardinal directions.

5 Conclusion

To recapitulate, Nganasan has a threefold category of number, which includes singular, dual, and plural. The number permeates the whole grammar and has to be obligatorily expressed in nouns and pronouns and, as agreement, in adjectives and verbs.

Virtually all nouns can be inflected for number. Plural forms can add different semantic effects, such as many instances of a substance for mass nouns (pieces of metal, clots of blood, reservoirs of water, etc.); a collective entity ('village' as a plural of 'house'); or can have an associative reading.

After numerals, the nouns are normally in the singular. On the other hand, the whole numeral phrase can be in the plural; then it either means an approximate number or refers to several groups of two, three, etc. entities.

There are no clear indications of the relevance of the number to the discourse structure; however, this question was never explored.

No verbal number exists, except for some cross-linguistically common aspectual derivations, such as iterative, frequentative, etc. No special verbs for singular or plural subjects or objects have been found so far.

Abbreviations

ABL	ablative
ABESS	abessive
ACC	accusative
ADJ	adjective marker
ALL	allative
ANT	anteriority
AUD	auditive
AUGM	augmentative
CARIT	caritive
CAUS	causative
CN	connegative
COMP	comparative ('like')
COND	conditional
CONNREC	connective-reciprocal
DEST	destinative
DIM	diminutive
DU	dual
DUB	dubitative
DUR	durative
FUT	future
EMPH	emphatic marker
EXCL	exclamative
GEN	genitive

IMP	imperative
IMPF	imperfective
INF	infinitive/converb
INFER	inferential
INTERJ	interjection
INTERR	interrogative
IPFV	imperfectivizer
ITER	iterative
HAB	habitual
LAT	lative
LIM	limitative ('only')
LOC	locative
MULT	multiplicative
NEG	negative auxiliary
NMLZ	nominalizer
O	objective conjugation
OP	objective conjugation for plural subject
PASS	passive
PF	perfect
PL	plural
PPF	pluperfect
PROL	prolative
PRES	present
PAST	past
PT	participle
PTCL	particle
PX	personal predicative marker
R	reflexive conjugation
RENARR	renarrative
RES	resultative
S	subjective conjugation
SG	singular
SUP	supine
VN.IPF	imperfective verbal noun
VN.PF	perfective verbal noun
VN.TEMP	temporal verbal noun
VOL	volitive

References

Brykina, Maria, Valentin Gusev, Sándor Szeverényi & Beáta Wagner-Nagy. 2018. *Nganasan Spoken Language Corpus (NSLC)*. Archived in Hamburger Zentrum für Sprachkorpora. Version 0.2. Available online at http://hdl.handle.net/11022/0000-0007-C6F2-8 (last accessed 08–01–2021).

Däbritz, Chris Lasse. 2017. Zur Markierung von Numerus an Nomina in den samojedischen Sprachen – Synchronie und mögliche Implikationen für die Rekonstruktion des Protosamojedischen und des Protouralischen. *Finno-Ugrische Mitteilungen*. 41. 63–101.

Daniel, Michael. 2005. The category of destinative in Nganasan, North-Samoyedic, and typology of prospective possession. Paper presented at the ALT VI in Padang, Sumatra, 21–25 July.

Daniel, Michael & Edith Moravcsik. 2013. The Associative Plural. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/36, Accessed on 2021 01 08.)

Gusev, Valentin. 2012. Aspekt v nganasanskom jazyke [Aspect in Nganasan]. *Acta Linguistica Petropolitana* VIII (2). 311–360.

Gusev, Valentin. 2020. Razve on dorogi ne znaet: K grammatikalizacii ritoricheskix voprosov s severosamodijskix jazykax [Doesn't it know the road? Grammaticalisation of rhetorical questions in North Samoyedic]. *Acta Linguistica Petropolitana* 16(3). 53–94.

Hajdú, Péter. 1975. *Samojedologische Schriften*. Szeged.

Helimski, Eugene [Khelimskij, E. A.]. 1994. Ocherk morfonologii i slovoizmenitel'noj morfologii nganasanskogo jazyka. In *Tajmyrskij etnolingvisticheskij sbornik. Vyp. 1*. 190–221. [Ethnolinguistic papers on Taimyr. Vol. 1]. Moscow: RGGU.

Helimski, Eugene. 1998. Nganasan. In Daniel Abondolo (ed.), *Uralic languages*, 480–515. London & New York: Routledge.

Helimski, Eugene. *Nganasan Dictionary*. Unpublished Manuscript.

Katzschmann, Michael 2008. *Chrestomathia Nganasanica. Texte – Übersetzung – Glossar – Grammatik*. Norderstedt: Books on Demand.

Khanina, Olesya & Andrey Shluinsky. 2014. A rare type of benefactive construction: Evidence from Enets. *Linguistics* 52. 1391–1431.

Leisiö, Larisa 2014. Imennye kategorii vremeni v nganasanskom i drugih severnosamodijskih jazykah. *Voprosy jazykoznanija* 1. 39–59.

Mikola, Tibor 2004. *Studien zur Geschichte der samojedischen Sprachen*. (Studia uralo-altaica 45). Szeged.

Nikolaeva, Irina. 2015. On the expression of TAM on nouns: Evidence from Tundra Nenets. *Lingua* 166. 99–126.

Plank, Frans 1989. On Humboldt on the Dual. In Roberta Corrigan, Fred Eckman & Michael Noonan (eds.), *Linguistic Categorization* (Amsterdam Studies in the Theory and History of Linguistic Science 61), 293–333. Amsterdam: John Benjamins.

Tereshchenko, Natal'ja M. 1979. *Nganasanskij jazyk* [Nganasan language]. Leningrad: Nauka.

Wagner-Nagy, Beáta. 2019. *A Grammar of Nganasan*. Leiden: Brill.

Ekaterina Gruzdeva

11 Number in Nivkh

Abstract: The chapter discusses the expression of number in Nivkh (isolate), which is spoken in the Far East of Russia. Typologically, Nivkh is a (poly)synthetic language with a rather complex (morpho)phonology and predominant suffixation. In Nivkh, number distinctions relate both to the count of participants, which is expressed by numerals and personal pronouns, as well as through nominal and verbal inflection, and to the count of events, which is determined at the clausal level, so that plurality of events is marked either by reduplication of nominal and verbal stems or by verbal suffixation. A maximum number of number features (singular vs. dual vs. plural) is attested in the personal pronouns and the hortative verb forms, whereas in all other forms a two-way number contrast between singular and plural is observed. The system of numerals contains at least thirty-three suffixalized classifiers. As an inflectional category, number is not obligatory either for nouns or for verbs. Those finite verb forms and converbs that can be conjugated display various patterns of verbal agreement. This can best be explained by assuming that the verbal forms are the results of morphosyntactic restructuring at different time levels.

1 Overview

Nivkh (Gilyak, Ghilyak) is an isolate language spoken on Sakhalin Island and in the Amur region of Russia. The language is traditionally classified as "Paleosiberian", or "Paleoasiatic", together with a number of other languages of Siberia and the Russian Far East. As a result of historical expansion from a homeland in central Manchuria to the Lower Amur and further to Sakhalin, Nivkh has been formed as a continuum of several distinct varieties – Amur (A), Liman (L), West Sakhalin (WS), North Sakhalin (NS), East Sakhalin (ES), Central Sakhalin (CS), and South Sakhalin (SS, extinct). Some of these varieties can be further divided into several subvarieties. For this paper, the most essential division is between the Tymovsk (ES^T) and Nyivo (ES^N) variants of the East Sakhalin variety. Amur and Liman Nivkh are spoken along the Lower Amur on the continent, while the other varieties are used on Sakhalin Island. A major split is observed between the Amur/Liman and East/Central/ South Sakhalin varieties, which actually fulfil the criteria for separate languages (the latter varieties are also known as Nighvng). The paper focuses essentially on the better studied Amur and East Sakhalin varieties, but when information from other varieties is available and relevant, it is also included in the discussion.[1]

---

1 All examples given in the paper without references were collected by the author on Sakhalin Island and in the Amur area between 1989 and 2019. The original transcription of examples taken from other sources is preserved.

Nivkh is a heavily endangered language: the current (2020) number of fluent Nivkh speakers is hardly more than 50, with fewer than 25 for the Amur, Liman and West Sakhalin varieties, only two for the North Sakhalin variety, and fewer than 25 for the Central and East Sakhalin varieties.

Typologically, Nivkh is a (poly)synthetic language with a rather complex phonology and predominant suffixation. Only possessor and object (including reflexive and reciprocal) markers can be prefixed (or more precisely pro-cliticized) to nouns and transitive verbs. Nivkh is well known by its root/morpheme-initial consonant alternations, which take place at morpheme boundaries and within noun and verb phrases.[2]

Nivkh noun morphology distinguishes two numbers (singular vs. plural), as well as eight cases in the Amur variety and seven cases in the East Sakhalin variety. Plural nouns are formed either by suffixation or by stem reduplication, or (rarely) by both of these means. Reduplication in plural formation is synchronically used only in the distributive context. All the nouns, except those denoting abstract concepts and homogeneous substances, can be pluralized. A special class is formed by singulative-collective nouns that display various models of singularization and pluralization. The nouns referring to kinship terms and proper names can form associative plural constructions by the regular plural suffix. Number contrast is also expressed by singular vs. plural vocative and comitative suffixes.

Suppletive forms of personal pronouns distinguish three persons (first, second and third), three numbers (singular, dual, and plural, with a further distinction between plural inclusive and plural exclusive) in the first person, and two numbers (singular and plural) in the second and third persons. Personal pronouns, as well as the reflexive-intensive pronoun A/ES *pʰi*, cannot attach the plural suffix, but can freely combine with case suffixes. Demonstrative, interrogative and indefinite pronouns, as well as the universal quantifier A *səkm*, ES *sikm* 'all (about humans)', A *sək*, ES *sik* 'all (about non-humans)' in the nominal function, inflect for number and case in the same way as nouns. Nivkh is also characterized by an elaborate system of suffixal numeral classifiers, which obligatorily occur in the numerals from 'one' through 'five' and optionally with some other numerals. No element of a numeral phrase is marked for number.

Quantitative aspectual meanings, such as multiplicative and distributive, are rendered by reduplication of nominal and verbal stems. Iterative plurality is marked by suffixation. In the indicative, finite verb forms optionally agree with the subject in number (singular vs. plural), but not in person. The verb has the whole paradigm of imperative, hortative and jussive forms, which is organized similarly to a personal pronominal paradigm with three numbers (singular vs. dual vs. plural) in the Amur variety and two numbers (singular vs. plural) in the East Sakhalin variety in

---

2 The words entering the complexes that are subject to morphophonological alternations are separated in the examples by '+'. For more details on Nivkh polysynthesis, see Mattissen (2003).

the first person (= hortative), two numbers (singular vs. plural) in the second person (= imperative) in all varieties, and two numbers (singular vs. plural) in the third person (= jussive) in the East Sakhalin variety. The syncretic differentiation of person and number (singular vs. plural) is found in coordinated, emphatic/evidential, negative assumptive, apprehensive and iterative forms.

At the level of the simple clause Nivkh displays some features of isolating structure, especially with respect to core arguments, which are typically unmarked. The canonical word order is *subject-object-verb* (SOV). All modifiers, except for numerals/quantifiers, precede the head nominal. Adverbial clauses are formed by converbs (= dependent forms) of different origin and are arranged in chains which end with the main clause. The latter, in turn, is built around the clause-final finite (= independent) form. Converbs render various interclausal relations, with the finest differentiation within the temporal ones. Some converbs agree with the subject in person and number similarly to aforementioned finite forms.

2 Pronominal, nominal, and verbal number

2.1 Generalities

In Nivkh, number distinctions relate both to the count of participants and to the count of events in which these participants take part.

The quantity of participants can be expressed by lexical items, in the first place by numerals and personal pronouns, as well as through nominal and verbal inflection. In the latter case the basic distinction is drawn between singular and plural forms. Singular forms are unmarked, whereas plural forms are created by suffixation. The dual number is restricted only to the first person personal pronouns and first person imperative forms. Depending on the form, a verbal predicate agrees with the subject either optionally or obligatorily. Part of numerals obligatorily comprise classifiers which are bound to the numeral elements.

The quantity of events is determined at the clausal level, so that plurality of events (multiplicative, distributive or iterative) is marked either by reduplication of nominal and verbal stems or by verbal suffixation.

The upcoming sections consecutively examine the marking of number in three lexical categories in the following order: pronominal number (2.2), nominal number, including numerals (2.3), and verbal number (2.4). Agreement of nominal and verbal forms is considered in section 3.

2.2 Pronominal number

In the class of personal pronouns, the number (together with the person) is expressed through a set of synchronically derived forms. Free forms of Nivkh personal

Tab. 1: Free forms of personal pronouns.

	A/L	WS	NS	ES/CS/SS
1SG	*ɲi*	*ɲi*	*ɲi*	*ɲi*
1DU	*megi, mege*	*meki, meke*	*memak*	*meŋ*
1PL INCLUSIVE	*mer, mir*	*mer, mir*	*mer, mir*	*mirn, miřn, min*
1PL EXCLUSIVE	*ɲəŋ*	*ɲəŋ*	*ɲəŋ*	*ɲin*
2SG	*cʰi*	*cʰi*	*cʰi*	*cʰi*
2PL	*cʰəŋ*	*cʰəŋ*	*cʰin*	*cʰin*
3SG	*if*	*if*	*i*	*jaŋ*
3PL	*imŋ, imx, ivŋ*	*imŋ*	*in*	*iřn, in*

pronouns are represented by eight suppletive forms that may differ depending on the variety and subvariety, cf. Table 1.

Dual number is attested only in the first person. Dual pronouns refer to a single speaker together with a single addressee. It seems that they can be used only with inclusive reference.

(1) **megi** *me-n*   *naf   tə + vəŋ + laqvt*   *u:mə-nə-ɟ=ra.*
     1DU   two-CL.HUMANS now this + cauldron + around fight-FUT-IND=FOC
     'Now we two will fight [running] around this cauldron.' [A]
     (Panfilov 1965: 228)

(2) **meŋ** *ŋatuk-ɣun=da.*
     1DU   person.of.the.same.age-PL=FOC
     'We two are of the same age.' [ES^T]

The dual pronouns are formally and semantically related to the numerals denoting 'two', namely to the A/L/WS/NS numeral *men* 'two', ES/CS/SS numeral *menŋ*, which is used for counting humans.[3] Both pronouns and numerals begin with the same numeral element **mi-* 'two' (see Section 2.3.6). As for the Amur and West Sakhalin dual pronouns *megi, mege, meki, meke*, their final components may be identified with the singular comitative suffix A/L/WS/NS *-Ke* (see Section 2.3.5).

In the first person plural a distinction is made between inclusive and exclusive reference. The inclusive pronoun refers to a speaker and non-speakers including the addressee(s):

---

3 More detailed information on the etymology of Nivkh personal pronouns can be found in Panfilov (1962: 204–214).

(3) *tʰət-ke* **mer** *mu ni-m-ɣir*
morning-CVB.DUR 1PL.INCL boat one-CL.BOATS-INSTR
u-ɣrə-t *vi-nə-ɟ.*
REC-be.with.sb.-CVB.NAR.1PL go-FUT-IND
'In the morning we shall go together by boat.' [A]
(Saveljeva 1960: 230)

(4) **min** *huŋ + kʰutə-roχ* *kuc-qavr-i-d=la?*
1PL.INCL that:CLOSE + hole-DAT fall-NEG-FUT-IND=Q
'Won't we fall into that hole?' [ESᴺ]

Austerlitz (1959: 109) suggests that initial *m* of the first person plural inclusive forms also derives historically from the numeral morpheme *⋆mi-* 'two', whereas the final part of the East Sakhalin and South Sakhalin pronouns *miřn, min* is connected with the third person plural pronoun *iřn, in*. If this is so, the formal structure of the first person plural inclusive pronoun perfectly reflects its reference to the speaker, addressee(s) and non-speakers.

First person exclusive pronouns refer to a speaker and non-speakers excluding the addressee(s):

(5) **ɲəŋ** *kʰezf-ujn* *niɲaq kʰrəu-ɟ.*
1PL.EXCL place.for.netting-LOC a.little rest-IND
'We took a rest in the place for netting.' [A]
(Otaina 1978: 97)

(6) **ɲin** *nud pʰ-řomsk* *ɣe-d-ɣun?*
1PL.EXCL what REFL-together take-IND-PL
'What did we take together [with] us?' [ESᵀ]

Nivkh singular personal pronouns also have bound forms that display allomorphy governed by the phonetic structure of the host element and regressive vowel harmony that entails the alternation of high vowels *i, ə, u* and non-high vowels *e, a, o*; cf. Table 2. Bound forms are employed when the pronoun (i) is case marked, cf. A/ES, *⋆ɲi-roχ > ɲe-rχ* <1SG-DAT> 'to me', *⋆cʰi-roχ > cʰe-rχ* <2SG-DAT> 'to you', *⋆i-roχ > e-rχ* <3SG-DAT> 'to him/her'; (ii) is used with relational nouns,[4] cf. A/ES *⋆ɲi-romsk > ɲ-romsk* 'together with me', *⋆cʰi-romsk > cʰ-řomsk* 'together with you', A *i-romsk,* ES *ja-romsk* 'together with him/her'; (iii) is used with discourse clitics, *⋆ɲi=an > ɲ=an* 'also me', *cʰi=an > cʰ=an* 'also you'; (iv) is attached to the head noun as a possessive marker, ES/A *cʰŋaj* 'image', *ɲe-zŋaj* 'my image', *cʰe-zŋaj* 'your image', A *ve-zŋaj* 'his/

---

4 Traditionally these morphological units are treated in Nivkh studies as postpositions.

her image'; (v) is attached to the head verb as an object marker, cf. ES/A *ɲ-za-* 'hit me', *cʰ-sa-* 'hit you', *i-ɟa-* 'hit him/her'. Bound forms display the same number values as the free forms.

Tab. 2: Bound forms of personal pronouns.

	A/L/WS/NS	ES	SS
1SG	*ɲi-* / *ɲe-* / *ɲ-*	*ɲi-* / *ɲe-* / *ɲ-*	*ɲi-* / *ɲe-* / *ɲ-*
2SG	*cʰi-* / *cʰe-* /*cʰ-*	*cʰi-* / *cʰe-* /*cʰ-*	*cʰi-* / *cʰe-* /*cʰ-*
3SG	*i-* / *vi-* / *e-* / *ve-* / *j-*	*ja-* / *ji-* / *je-* / *i-* / *e-* / *j-*	*ji-* / *je-* / *j-*

Neither the singular nor the dual personal pronouns can attach the plural suffix in any Nivkh variety. As for the Amur plural personal pronouns, according to Panfilov (1962: 235) they can occasionally attach the regular plural suffix *-ku* (see Section 2.3.1):

(7) 1PL *ɲəŋ > ɲəŋ-gu* 'we'
     2PL *cʰəŋ > cʰəŋ-gu* 'you'
     3PL *imŋ > imŋ-gu* 'they' [A]

Otaina (1960: 228) states that such pluralized pronominal forms are found mostly in the spoken language and are used for emphasis. This is basically confirmed by my data as well. In (8a) the pronominal subject *imŋ* 'they' does not have any special prominence, so the plural suffix is attached to the number-agreeing finite verb. By contrast, in (8b) the pronominal subject is foregrounded and therefore takes the plural suffix, while the finite verb is left unmarked for plural.

(8) a. **imŋ** *pʰrə-ɟ-ɣu.*
        3PL come-IND-PL
        'They came.' [A]

     b. **imŋ-gu** *pʰrə-ɟ.*
        3PL-PL come-IND
        'They (exactly) came.' [A]

Personal pronouns may be part of a comitative construction and thereby attach singular and plural comitative suffixes (see Section 2.3.5). Singular personal pronouns occur only with the singular comitative suffix as shown in Table 3:

Tab. 3: Comitative (single) forms of the singular personal pronouns.[5]

	A/L/WS	ES^N	ES^T	
1SG	ɲi-ɣe	ɲi-ɣin	ɲi-ɣen	'with me'
2SG	cʰi-ɣe	cʰi-ɣin	cʰi-ɣen	'with you'
3SG	if-ke	jaŋ-gin	jaŋ-gen, i-ɣen	'with him'

Tab. 4: Comitative (plural) forms of the plural personal pronouns.

	A/ L/WS	ES	
1PL	ɲəŋ-go(n)	ɲin-ɣunu	'with us'
2PL	cʰəŋ-go(n)	cʰin-ɣunu	'with you'
3PL	imŋ-go(n)	in-ɣunu	'with them'

Consider the following example:

(9) **ɲi-ɣen**   pʰ-ŋafq-ɣen   ŋatu-d-ɣun.
    1SG-COM.SG REFL-friend-COM.SG be.of.the.same.age-IND-PL
    'Me and my friend are of the same age.' [ES^T]

Plural pronouns in turn occur only with the plural comitative suffix, as shown in Table 4.

The examples are:

(10) nama-gur   kʰəmlə-rot   **imŋ-gon**   va-ja.
     be.good-ADV.2SG think-CVB.DIST.2SG 3PL-COM.PL fight-IMP.2SG
     'After thinking well, fight with them.' [A]
     (Panfilov 1965: 224)

(11) hana **in-ɣunu**   lu-da.
      let   3PL-COM.PL sing-HORT.1PL
      'Let us sing with them.' [ES^N]

Demonstrative, interrogative and indefinite pronouns, as well as the universal quantifier A *səkm*, ES *sikm* 'all (about humans)', A *sək*, ES *sik* 'all (about non-humans)' in the nominal function can freely attach the plural suffix A -*Ku*, ES -*Kun* cf.

---

5 The allomorphy of the comitative suffixes basically follows the same rules as that of the plural suffix (see Section 2.3.1). However, there is some irregular variation when they are used with personal pronouns (not discussed in this article).

ES *hud-ɣun řawd-ɣun=ara?* <that.CLOSE-PL who-PL=Q> 'Who are those ones?'; A *sək-ɣu + pʰur-ɣət-c* <all-PL + tell-COMPL-IND> '[s/he] told everything'. In my data there are no examples of the use of comitative suffixes with any pro-forms, besides personal pronouns.

2.3 Nominal number

This section deals, first, with the nominal number inflection which is synchronically rendered by suffixation (2.3.1). Second, it considers the semantic classes of nouns that are relevant for a description of number marking (2.3.2). Third, it explains how nominal reduplication is used for expressing distributivity and iterativity (2.3.3). Then it examines the marking of number in the vocative (2.3.4) and comitative forms (2.3.5). Finally, it discusses numerals and numeral classifiers (2.3.6).

2.3.1 Number inflection

Nivkh nominal morphology distinguishes two numbers: an unmarked singular that is opposed to a plural formed by the suffix A/L/WS -*Ku,* NS/ES/CS/SS -*Kun.* The suffix is attached to the nominal stem and can be followed by a case suffix, by a relational noun with or without a case suffix, or by a word-final discourse clitic, cf. A *təf* <house:SG.NOM> 'house', *təf-ku-doχ* <house-PL-DAT> 'to (the) houses', *təf-ku + mi-ujn=lu* <house-PL + inside-LOC=PROB> 'probably in the houses'. The allomorphic variation of the plural suffix follows the general rules of Nivkh morphophonological alternations, which are based on two historical processes: spirantization of stops in the intervocalic context and sonorization of weak stops in the post-nasal context. The choice of the initial consonant of the suffix depends on the final sound of the stem to which it is attached:[6]

(i) A -*ku,* ES -*kun* is used after stems ending in fricatives (including voiceless trill *ř*): A *təf-ku* 'houses', ES *hajmŋař-kun* 'old men';

(ii) A -*ɣu,* ES -*ɣun* is used after stems ending in a vowel, a lateral, a stable nasal,[7] the approximant *j* and stops: A *mu-ɣu* 'boats', ES *tʰlaŋi-ɣun* 'reindeers';

(iii) A -*gu,* ES -*gun* is used after stems ending in a non-stable nasal: A **qanŋ-gu > qan-gu* 'dogs', ES *eɣlŋ-gun*[8] 'children'.

---

6 Note that the transcription of plural (and other) forms depends on the author and may deviate from these rules. This is especially evident for the post-nasal context and can be a reflection of ongoing variation and change.

7 Stable nasals have been historically retained in all Nivkh varieties, whereas non-stable nasals have been lost in the A, L and WS varieties and retained in other varieties (see e.g. Kreinovich 1937; Panfilov 1962; Gruzdeva 1997a).

8 In the modern ES variety, the nasal *ŋ* is often lost in the final segment of noun stems before a number suffix. This phenomenon is connected with the general diachronic process of nasal loss.

Plural number is not obligatorily marked on the nouns. Only a prototypical count noun which refers to a discrete, individualized referent and appears in the sentence as a subject regularly shows the singular-plural distinction. The presence of the plural marker also depends on the Animacy Hierarchy, since nouns denoting animates are more likely to be marked for number than inanimates:

(12) *ɲəŋ-doχ* **antx-ku** *pʰrə-ɟ-ɣu.*
1PL.EXCL-DAT guest-PL come-IND-PL
'The guests came to us.' [A]
(Saveljeva and Taksami 1970: 527)

(13) *tʰulf* *ha-data* **əcx-kun** *qʰavla + tʰolf* *ha-jnə-vut*
winter do.so-CVB.BSIM old.man-PL be.hot + summer do.so-DES-QUOT.3PL
it-nt-ɣun.
say-IND-PL
'During the winter old men said that the summer was going to be hot.' [ES]

Moreover, even a noun meeting the above-mentioned conditions may occur in the singular if the plural suffix is taken by the verbal predicate:

(14) *imŋ + * **qan** *əɣ-ɟ-ɣu.*
3PL + dog bark-IND-PL
'Their dogs barked.' [A]
(Panfilov 1962: 99)

The general tendency is to mark plurality once – either on the subject or on the verbal predicate (see Section 3.3 on the syntax of number). Thus, subjects tend not to be marked for plural when used with predicates denoting an action that is simultaneously performed by several participants. This is illustrated by (15), where the verb *maɣmaɣ-* 'to swim up in a shoal, in a group' implicitly points to a plural reference of the subject *laŋr* 'seal':

(15) *ɟ-ordoχ* *pila + baχ-ku-ux* **laŋr** *maɣ~maɣ-ɟ=ra.*
3SG-towards be.big + stone-PL-ABL seal swip.up.in.a.shoal~RED-IND=FOC
'Seals swam up in a shoal towards him from behind big stones.' [A]
(Saveljeva and Taksami 1970: 172)

Also, in the object function, nouns with a plural reference are not necessarily marked for plural, cf. the nouns ES *va* 'sabers', *qʰaχ* 'spears', *puɲd* 'bows', *ku* 'axes', and *cʰχař* 'sticks' in (16). However, in a similar context they can also take a plural suffix cf. *pʰ-fund-ɣun=lo* 'bows' and *pʰ-χaχ-kun=lo* 'spears', as in (17). One may hy-

pothesize that in a command in (17) the nouns in the object function are more emphasized than those in the narrative sentence in (16).

(16) *hə + ɲiɣvŋ-gun*   ***va + řo-nd.***   ***va + řor-ra***
      that:CLOSE + man-PL saber + carry-IND saber + carry-COORD.3SG
      qʰaχ + tʰo-nd.   ***qʰaχ + tʰor-ra***   ***puɳd + řo-nd.***
      spear + carry-IND spear + carry-COORD.3SG bow + carry-IND
      řor-ra   ***ku + řo-nd.***   *řor-ra*   ***cʰχař + tʰo-nd.***
      carry-COORD.3SG axe + carry-IND carry-COORD.3SG tree + carry-IND
      tʰor-ra.
      carry-COORD.3SG
      'These men held sabers. [He] held sabers, held spears. [He] held bows.
      [He] held axes. [He] held sticks.' [ES]
      (Shternberg 1908: 105)

(17) *ol-ɣun-a*   ***pʰ-fund-ɣun=lo***   ***pʰ-χaχ-kun=lo + ɣe-ve.***
      child-PL-VOC REFL-bow-PL=PROB REFL-spear-PL=PROB + take-IMP.2PL
      'Children, take your bows [and] spears!' [ES]
      (Shternberg 1908: 45)

Irrespective of syntactic function, a count noun is marked for plural when it is used with the universal quantifier A *səkm*, ES *sikm* 'all (about humans)', A *sək*, ES *sik* 'all (about non-humans)', cf. (18)–(19). The quantifier A/ES *les* 'many, enough' seems to be less strict in this respect, but in the modern language is typically also used with the quantified noun marked for plural, including with mass nouns in (30) and (72).

(18) ***i-marq-xu*** *sək* *tol-uin*   *huŋɲi-tʰa-ve.*
      3SG-bone-PL all   water-LOC leave.there-ITER.2PL-IMP.2PL
      'Leave all her bones near the water.' [A]
      (Panfilov 1965: 249)

(19) *ɲe-rχ*   ***pʰ-χalŋ-gun***   *sikm* *xeř + pʰur-ja.*
      1SG-DAT REFL-relative-PL all   news + tell-IMP.2SG
      'Tell me the news [about] all my relatives.' [ES]

This quantificational meaning can also be expressed adverbally with the intransitive verbs A/ES *malʁo-* 'be abundant' and A *həsk-*, ES *hasku-* 'be small (in quantity)', cf. (21), or with the transitive verb A/ES *ram-* 'be abundant [in sth.]' in the predicative function. Each of these verbs can be used both in the predicative and attributive functions, and none of them requires a quantified argument in the plural form, cf. (20)–(23). However, in the speech of modern speakers the pluralization is very frequent, cf. (24)–(25).

(20) *jaŋ* **eɣlŋ malʁo-d**.
 3SG child be.abundant-IND
 'She [has] many children.' [ES^T]

(21) *ɲi* **malɣo-gut** **mař-kiř** *pəiz-ŋan* *c^həŋ-rox*
 1SG be.abundant-ADV.1SG scale-INSTR throw-CVB.SIM 2PL-DAT
 malɣo-la + **c^ho** *vi-řa-ɟ=ra.* *ɲi* **həsk-kut**
 be.abundant-QUAL + fish go-ITER.2SG-IND=FOC 1SG be.small-ADV.1SG
 mař-kiř *pəiz-ŋan* **c^ho həsk-kuř** *vi-řa-ɟ=ra.*
 scale-INSTR throw-CVB.SIM fish be.small-ADV.3SG go-ITER.3SG-IND=FOC
 'If I throw much scale, a lot of fish come to you. If I throw little scale, little
 fish comes.' [WS]
 (Panfilov 1965: 245)

(22) *t^hvaχ* **ŋa + ram-d**.
 small.gulf seal/beast + be.abundant-IND
 'The gulf is abundant in seals.' [ES^T]

(23) *hoʁoŋan məxckiř əlɣi-la + vo* **c^ho + ram-ɟ-roɣo**
 then really be.rich-QUAL + settlement fish + be.abundant-NMLZ-DEST
 laŋř + tam-ɟ-roɣo.
 seal + be.abundant-NMLZ-DEST
 'Then [there is] really a rich settlement abundant in fish, abundant in seals.' [WS]
 (Panfilov 1965: 242)

(24) *hemar* **malʁo-la + ŋəzit-ɣu + him-ɟ**.
 old.man be.abundant-QUAL + tale-PL + know-IND
 'The old man knows many tales.' [WS]

(25) *huŋ + vo-wx⁹* **ɲiɣvŋ-gun hasku-t** *jiv-d-ɣun.*
 that:CLOSE + settlement-ABL man-PL be.small-CVB.NAR.3PL exist-IND-PL
 'Few people live in this settlement.' [ES^T]

Another way of coding plurality is by reduplicating a noun stem. Panfilov (1962:
96) suggests that originally reduplication has been the most widespread device for
marking the plural number in Nivkh. This claim can be supported by some exam-
ples attested in the folklore texts, cf. A *ciɣř~ciɣř* 'trees' in (26).

---

9 The ablative case can be used for expressing both location and source.

(26) *ivŋ + təf + tul-ɣu=pařk*   *pal-ux*   *hum-ta*   **ciɣř~ciɣř**
   3PL + house + hole-PL=only forest-ABL be-COORD.3PL tree~RED
   pant-ɣər-ta ...   *vax* *pant-ɣət-ra.*
   grow-COMPL-COORD.3PL moss grow-COMPL-COORD.3SG
   'There are only holes of their houses in the forest, trees have grown up, moss
   has grown up.' [A]
   (Panfilov 1965: 242)

2.3.2 Plural across special lexical classes of nouns

From a semantic point of view, nouns are divided into count nouns with a regular
and productive number inflection (see Section 2.3.1), *singularia tantum* abstract
nouns, mass nouns that can be pluralized only when referring to specific types of
substances, and several types of (singulative)-collective nouns, most of which ex-
press both an individual object and a set of objects. There are no *pluralia tantum*
nouns.

Plural cannot be marked on abstract nouns denoting qualities, actions, states,
or general ideas, such as the following (taken, in particular, from Panfilov 1962:
116–117; Saveljeva and Taksami 1970; Sangi and Gashilova 2003): A *uls*, ES *ulř*
'height', A *kəls*, ES *kəlř* 'length', A *civla*, ES *civlaŋ* 'cold',[10] A *qʰavla*, ES *qʰavlad*
'heat', A *cʰolaɟ*, ES *cʰolad* 'poverty', A *qʰorlaɟ* 'wealth', A *umlaɟ* 'malice', A *kʰərx*
'illness', A *məks*, ES *maxtuř* 'truth', A *humɟ*, ES *hunmf* 'existence', A *rajuɟ* 'writing',
A *juruɟ* 'reading', A *valtc* 'lie', A *ezmuɟ* 'gladness', etc. In the Amur variety, a large
number of abstract nouns are lexical nominalizations derived from the stems of
qualitative verbs with the suffix *-ɟ/-c*. In general, abstract nouns are used in narra-
tives extremely rarely and appear mostly in elicited examples.

2.3.2.1 Mass nouns
Mass nouns indicating homogeneous substances or substances with a minimal unit
normally do not take a plural suffix, cf. (27). Examples are A/ES *cʰaχ* 'water', A/ES
cʰoχ 'juice, resin, infusion, blood', A *pəŋx*, ES *pɑŋx* 'broth', A/ES *mos* 'fish-jelly',
A/ES *tʰom* 'grease', A/ES *ɲoχ* 'fat', A *cus*, ES *tuř* 'meat', A/ES *lep* 'bread', A/ES *raq*
'groats, gruel', A *ova*, ES^N *ovaŋ*, ES^T *ofaŋ* 'flour', A *tafc*, ES *tafciŋ* 'salt', A/ES *seta*
'sugar', A/ES *tamx* 'tobacco', A/ES *oχt* 'powder, cure, medication', A/ES *tʰeŋ* 'coal',
A *vəc*, ES *vat* 'iron', A *ajs*, ES *ajzŋ* 'gold', A/ES *pos* 'cloth', A *lur*, ES *luř* 'ice', etc.
However, if a mass noun comes to refer to a specific type of mass entity, it can be
pluralized, as the noun *pos* 'cloth' in (28):

---

10 Panfilov (1962: 116) points out that in plural this word has another meaning, i.e. *civla-gu* 'frosty
mist'.

(27) *siŋ~siŋ + **iɲnəɟ** malɣo-ɟ.*
what~RED + food be.abundant-IND
'There are many different foods.'[A]
(Panfilov 1965: 250)

(28) *maɟu-ɣun poci~voci-ŋ + **bos-kun-gis** veřki-d-ɣun.*
Chinese-PL be.various~RED-NMLZ + cloth-PL-INSTR sell-IND-PL
'Chinamen sell various types of clothes.' [ES^T]

Similar cases of pluralization of several mass nouns, i.e., *raq* 'groats', *ova* 'flour', *tafc* 'salt', *tʰom* 'grease', are recorded in the Amur example (29), which allows two possible readings. The plural forms of mass nouns may have either a qualitative interpretation, i.e., designate different kinds of groats, flour and grease, or a quantitative one, i.e., denote several entities of groats, flour, salt and grease:

(29) *hə + tʰuɣrmu **raq-xu ova-gu tafc-ɣu***
that:CLOSE + steamer groats-PL flour-PL salt-PL
***tʰom-gu** + tʰo-r pʰrə-ɟ.*
grease-PL + carry-CVB.NAR.3SG come-IND
'This steamer came bringing groats [= 'several kinds or many sacks'],
flour [= 'several kinds or many sacks'], salt [= 'many sacks'],
grease [= 'several kinds or many vessels'].' [A]
(Panfilov 1962: 110)

My modern East Sakhalin consultants consider such examples in their own variety as incorrect, pointing out that of the four mass nouns used in (30), only *raq* 'groats' can be pluralized and even then only the reading 'several kinds of groats' is possible:

(30) *ɲi tu + ɲo + mi-x poci~voci-ŋ + **raq-ɣun** les si-d.*
1SG this + store + inside-ABL be.various~RED-NMLZ + groats-PL many put-IND
'I put many different kinds of groats into the store.' [ES]

In an appropriate context some mass nouns acquire the meaning of countable nouns, cf. *ajs-ku* 'golds' that refers to different things made of gold:

(31) *if řaŋga + **ajs-ku** jiv-ɟ.*
3SG many + gold-PL exist-IND
'He has a lot of gold [= 'golden things'].' [A]
(Panfilov 1962: 110)

Furthermore, Panfilov (1962: 111) argues that the plural form of some mass nouns (apparently in combination with the universal quantifier) may indicate a great

amount of mass entity, as in (32), where *cus-ku* denotes '[all] meat of the hero's body':

(32) **i-cus-ku**   *sək vəc + mu-ɣət-c.*
     3SG-meat-PL all iron + become-COMPL-IND
     'All his meat turned into iron.' [A]
     (Panfilov 1965: 229)

Once again, such a reading seems impossible to the modern speakers of the East Sakhalin variety. In the similar example (33), the plural form *tuř-kun* is interpreted only as 'various kinds of meat':

(33) *ɲin   sik pʰi   iv-ŋ + **tuř-kun** + ɲi-χar-d-ɣun.*[11]
     1PL.EXCL all REFL exist-NMLZ + meat-PL + eat-COMPL-IND-PL
     'We have eaten all kinds of meat we had.' [ES^T]

2.3.2.2 Collective nouns

A separate semantic class is represented by singulative-collective nouns that can be further subdivided into three classes.[12]

Singulative-collective nouns from the first class denote either: (i) a single object (often of small size) or (ii) a homogeneous set or substance composed of corresponding objects, cf. A *həɣr,* ES *ŋawk* 'grain of roe / roe', A *ŋaqr,* ES *qʰavi* 'snow-flake / snow', ES *ŋavřki* 'hair (one) / hair (of an animal)', A *ɲevrqaj,* ES *ɲevřki* 'eyelash / eyelashes', A/ES *kʰu* 'pellet / small shot', A *alʁas,* ES *alʁař* 'bead / beads', A *ŋəŋg,* ES *ŋamx* 'hair (one) / hair (of a man)', A *max,* ES *qomř* 'grain of sand / sand', A *cʰŋər,* ES *cʰŋəř* 'blade of grass / grass', A *als,* ES *alř* 'berry / berries', A *comr,* ES *comř* 'leaf / leaves', A *tʰoɣs,* ES *tʰoɣzŋ* 'cone / cones', A *pus,* ES *puř* 'speck of dust / dust', A *ŋəks,* ES *ŋakř* 'branch / bushes'. A subgroup of this class is represented by nouns referring to a plural body part / parts other than pairs, cf. A *ŋəɣs,* ES *ŋaɣzəř* 'tooth / teeth', A/ES *ems* 'molar / molars', A *tuɲ,* ES *tuɲm* 'finger / fingers', A *təkɲ,* ES *takɲ* 'nail / nails', A/ES *ŋarm* 'rib / ribs'. As can be seen from the following examples, the noun *qomř* can have either a singulative meaning of 'a single grain of sand', cf. the elicited example (34), or a collective meaning of 'sand', cf. (35):

(34) *ɲ-ɲaʁ-ux   **qomř** + kuz-ja.*
     1SG-eye-ABL sand + take.out-IMP.2SG
     'Take [a grain of] sand out of my eye!' [ES^T]

---

11 In the ES variety, the verb *jiv-* glossed as 'exist' can also be used as a transitive verb with the meaning 'have'.

12 For more discussion and data from the Amur variety, see Panfilov (1962: 101–111).

(35) *eɣlŋ-gun* **qomr-ux** *ler-d-ɣun.*
child-PL sand-ABL play-IND-PL
'Children play in the sand.' [ES^T]

All nouns of this class can be regularly pluralized. The resulting form is typically interpreted as 'several sets of similar objects'. This can be illustrated by (36), where *qomř-kun* denotes different kinds of sand. However, even in this sentence a singular form *qomř* is admissible.

(36) *i + orr-ux   pʰfala~vala-ŋ +* **qomř** / **qomř-kun**   *jiv-d.*
river + bank be.different~RED-NMLZ + sand / sand-PL exist-IND
'There are different [kinds] of sand on the bank of the river.' [ES^T]

Without a context the meaning of these nouns is typically interpreted as collective. Whenever necessary, the singulative meaning can be actualized by using the numeral *ɲi-k* <one-CL.THREE.DIM.OBJ> 'one [referring to three-dimensional (small) objects]', cf. A *als ɲik* 'one berry', ES *ŋawk ɲik* 'one [grain of] roe', ES *alʁař ɲik* 'one bead':

(37) *ŋawk* **ɲi-k** *+ ɣe-ř   amam-d.*
roe   one-CL.THREE.DIM.OBJ + take-CVB.NAR.3SG taste-IND
'Taking a grain of roe, [he] tasted [it].' [ES^N]

Another way to singularize such nouns is to use a lexical singulative noun, such as *xem* 'grain', which can be followed by the numeral *ɲik*:

(38) *ŋafq-a   ŋafq-a   **həɣr + xem ɲi-k**=hara*
friend-VOC.SG friend-VOC.SG roe + grain one-CL.THREE.DIM.OBJ=COORD
kəp + xem   ɲi-k=hara   *ɲ-əl + fəɲ-ja.*
bird.cherry + grain one-CL.THREE.DIM.OBJ=COORD 1SG-mouth + throw-IMP.2SG
'Friend, friend, throw a grain of roe, a grain of bird cherry [into] my mouth.'
[A]
(Kreinovich 1934: 221)

One of the most peculiar singulative-collective nouns in Nivkh is A *ciɣr / ci:r*, ES *cʰχař*, which is characterized by an exceptionally wide range of meanings, cf. 'tree' in (39), 'trees' in (40), 'stick' in (41), 'firewood' in (42), and 'forest' in (43):

(39) *la   kuɣu-ŋ +* **cʰχař** *+ volu-j-d.*
wind stand-NMLZ + tree + fell-FUT-IND
'The wind will bring a standing tree down.' [ES]

(40) *hor   j-ama-ŋan   lil tol-kar   tol-kar   kəl-kar*
then:3SG 3SG-look-CVB.SIM even be.thick-AUG be.thick-AUG be.long-AUG
kəl-kar + ci:r + moq~moq-r   məγ-ɟ.
be.long-AUG + tree + cut~RED-CVB.NAR.3SG come.down-IND
'Then when [he] looked, [a bear] breaking even very thick [and] very long trees came down.' [A]
(Otaina 1978: 58)

(41) *ɲi   cʰχař-kis   kʰeq + xutə + γuř~kur-d.*
1SG tree-INSTR fox + hole + grope~RED-IND
'I groped with a stick in a fox's hole.' [ESᵀ]

(42) *tʰuγr-ux cʰχař teγ~teγ-d.*
fire-ABL tree   crack~RED-IND
'The firewood cracked in the fire.' [ES]

(43) *vo + əx-tox   mər-ŋan   ci:r + mi-rx   mər-ɟ.*
village + end-DAT come.up-CVB.SIM tree + inside-DAT come.up-IND
'When he came up to the end of the village, he came up into a forest.' [A]
(Panfilov 1962: 108)

In the plural A *ciγr* / *ci:r*, ES *cʰχař* refers only to 'trees', cf. (44), 'sticks' or 'firewood', cf. (45):

(44) *pʰxi-wx   pol-ŋ + cʰχař-kun + tʰmə~řmə-t   vi-vul*
forest-ABL fall-NMLZ + tree-PL + climb.over~RED-CVB.NAR.1PL go-CVB.UBSIM
eγ-gut   pʰer-d-γun.
quickly-ADV.1PL be.tired-IND-PL
'When [we] were going in the forest climbing over the fallen trees, [we] quickly became tired.' [ES]

(45) *hoʁorot cəj   i-cir-ku-gir   malgo-gu-r*
then:3SG again 3SG-tree-PL-INSTR be.abundant-CAUS-CVB.NAR.3SG
tʰu:r + pʰu-ɟ.
fire + make-IND
'Then [he] again made a big fire of his firewood.' [A] (Panfilov 1962: 100)

The second class of singulative-collective nouns comprises lexical items denoting either (i) one of a pair (including paired body parts) or (ii) a pair as a whole, cf. A/ES *tot* 'arm / arms', A *təmk,* ES *tamk* 'hand / hands', A *cʰəlm,* ES *cʰalm* 'palm / palms', A *ŋəcx,* ES *ŋacx* 'leg / legs', A *ŋəzl,* ES *ŋazl* 'foot / feet', A/ES *ɲaχ* 'eye / eyes', A *nos*, ES *mla* 'ear / ears', A *vixkʰutə,* ES *uxkʰutə* 'nostril / nostrils', A *ŋəɲk,*

ES *ɲetf* 'cheek / cheeks', A *itə*, ES *heɣmi* 'jaw / jaws', A *tʰəws*, ES *qařqŋalχ* 'kidney / kidneys', A *ŋərɣər,* ES *ŋajrəx* 'wing / wings', A *ŋəsq,* ES *ŋasq* 'seal's back flipper / flippers', A *murki*, ES *muřki* 'horn / horns', A/ES *vamq* 'mitten / mittens', A/ES *mesq* 'ear ring / ear rings', A *ujɣir,* ES *ujɣiř* 'sleeve / sleeves', A *əvɲ*, ES *kikř* 'oar / oars', A/ES *mizurəx* 'rowlock / rowlocks', A/ES *ki* 'boot / footwear', A/ES *pʰaɲ* 'shin-pad / shin-pads', A/ES *laq* 'ski / skis', A/ES *eɲ* 'fur ski / fur skis', etc. The following examples show that the noun *ɲaχ* may refer both to one eye, as in (46), and to eyes, as in (47):

(46) *əcx   ɲřaklu   tʰajrə-ŋ + **ɲaχ-kis**   ɲel~ɲel-d.*
old.man sometimes be.left-NMLZ + eye-INSTR wink~RED-IND
'The old man sometimes winks with [his] left eye.' [ES^T]

(47) *ɲ-ziɣ-ŋ + řaŋɢeɣlŋ-ux   **ɲaχ** vəl~vəl-d.*
1SG-arrange-NMLZ + girl-ABL eye be.dark~RED-IND
'My bride [has] dark eyes.' [ES^T]

Historically nouns designating paired objects occurred predominantly in the singular (Panfilov 1962: 107), taking the plural suffix only when it was necessary to stress their non-singularity. This can be seen from (48), where the storyteller emphasizes that the hero of the story rubbed grease into his two legs (*ŋəcx-ku*) and arms (*tot-ɣu)*. Here, the usage of the plural forms is also related to the utilization of the universal quantifier A *sək* 'all'.

(48) *hə + tʰom-gir   sək **ŋəcx-ku tot-ɣu** + xedr-ɟ.*
that:CLOSE + grease-INSTR all leg-PL   arm-PL + rub-IND
'[He] rubbed this grease into all legs and arms.' [A]
(Panfilov 1962: 107)

When pluralized, nouns referring to such paired objects usually denote several pairs of them, cf. (49), where *ɲaχ-ku* 'eyes' refer to the eyes of several wolves:

(49) *plaju + **ɲaχ-ku** eʙ-gur   ma-jo-ɟ.*
flash + eye-PL   quickly-ADV.3SG be.close-DIM-IND
'Flashing eyes are quickly approaching.' [A]
(Panfilov 1962: 108)

However, in modern Nivkh there is a clear tendency to use the plural suffix when reference is made to the two members of the pair. This tendency is probably caused by the influence of Russian, where in such cases the plural form is required. For instance, in two consecutive sentences, cf. (50), a bilingual consultant used two different forms for expressing the same meaning 'hands'. In the first sentence the

noun *tamk-ɣun* is marked for plural (as in Russian), whereas in the second one the traditional singular form *tamk* is used:

(50) **pʰ-řamk-ɣun** + ɲřə-l? **tamk** eɣřk-t.
  REFL-hand-PL + see-Q hand be.dirty-IND
  'Did [you] see your hands? The hands are dirty.' [ES]

The standard (but gradually lost by the modern speakers) way to singularize nouns from this class is by using the noun A *pasq*, ES *pazř* 'half', cf. A *ujɣir + pasq* '[one] sleeve', ES *tamk + vazř* '[one] hand', A *əvŋ + pasq* '[one] oar', ES *vamq + vazř* '[one] mitten', and (51):

(51) **tot** + **vazř**=vəřk-ŋ   + atkəcx
  arm + half=only-NMLZ + old.man
  'an old man with only one arm' [ES]

Nouns denoting objects consisting of two inseparable parts, like A *χaza*, ES *χazaŋ* 'scissors', A/ES *vac* 'tongs', A/ES *var* 'trousers', etc., are normally singular. Like count nouns they attach the plural suffix only when referring to several objects (several pairs of scissors, etc.).

The third (rather limited) class of singulative-collective nouns consists of lexical items denoting either (i) a single object as a part of a heterogeneous class of objects used for a similar purpose or (ii) a heterogeneous class of such objects, cf. A *aʁaɣs*, ES *aʁaɣř* 'thing / things', A *ŋaŋənaʁaɣs*, ES *ŋaŋanɣŋaʁaɣř* 'a piece of hunting equipment / hunting equipment', A *cʰoŋənaʁaɣs,* ES *cʰoŋanɣŋaʁaɣř* 'a piece of fishing equipment / fishing equipment', A/ES *oq* 'a piece of outer clothing / outer clothing', A *larq*, ES *lařq* 'a dress, a piece of underclothes / underclothes', A *uŋs,* ES *uŋř* 'a piece of crockery / crockery', A *ŋir*, ES *ŋirŋ* 'a dish / dishes', A *laqrs*, ES *laqř* 'a toy / toys', A *ves*, ES *veř* 'sth. of goods / goods'.

As in other cases, the noun *oq* 'outer clothing' can, in principle, indicate either a single article of clothes or a collection of clothes, cf. (52). However, for this group of nouns a singulative meaning should be considered a prototypical one.

(52) *acm*   **oq** + *ot-t.*
  grandmother outer.clothing + sew-IND
  'Grandmother sews the outer clothing(s).' [ES]

Furthermore, referring to a set of objects, the nouns from this class are quite often marked as plurals, especially when heterogeneity of objects is obvious from the context, cf. *oq-ɣun* with the meaning 'a collection of different clothes' in (53), and 'collections of clothes made by different people' in (54):

(53) *magazin-doχ naf cʰir-ŋ + **oq-ɣun** + tʰo-t*
     shop-DAT   now be.new-NMLZ + clothes-PL + carry-CVB.NAR.3PL
     pʰy̆ə-d-ɣun.
     come-IND-PL
     'Now [they] bring to the shop new clothes.' [ES]

(54) *pocur-la-ŋ + **oq-ɣun** + ot-t-ɣun.*
     be.beautiful-QUAL-NMLZ + clothes-PL + sew-IND-PL
     '[They] sew nice clothes.' [ES]

A separate group is formed by collective nouns that can only denote a set of elements, cf. A *iks,* ES *ik̆r* 'team of nine dogs', A *nuxc* 'team of nine dogs', ES *lele* 'relatives',[13] A *ruv,* ES *ruvŋ* 'family clan', A *qalaɟ,* ES *təjk* 'greens', A *uzi,* ES *uzu* 'firewood'. Being marked for plural, some of these nouns refer to abundance of quality, similarly to mass nouns, while the others refer to several sets of objects. Thus, in (55) the meaning of the plural noun *lele-gu* 'relatives' may be interpreted either as 'all relatives' or as 'different relatives':

(55) *if   **pʰ-lele-gu-doχ**   kʰinŋu-ɟ.*
     3SG REFL-relatives-PL-DAT be.nice-IND
     'He respects his relatives.' [A]
     (Panfilov 1962: 108)

2.3.2.3 Associative plural

Plural forms of kinship terms and proper names may render associative plural meaning, i.e. refer to a person together with other people in some way connected with him/her (Panfilov 1962: 113–114). The typical examples are ES *ətk-ɣun* 'father and his family', ES *əmɣi-ɣun* 'son-in-law and his family', ES *aŋɣej-ɣun* 'wife and her relatives', cf. *pʰ-əmk-ɣun-gis* 'about his mother and her family' in (56) and *Xevgun-gu* 'Xevgun together with other people' in (57). Modern speakers use such forms very rarely.

(56) *eɣlŋ ŋawzŏr-fke   pʰaki~faki   tʰəxt-nd   **pʰ-əmk-ɣun-gis***
     child grieve-CVB.DUR gradually~RED be.amused-IND REFL-mother-PL-INSTR
     tʰχăr̆p-ifu-nd=ajaq.
     forget-PROGR-IND=MOD
     'A child after grieving for a long time, little by little starts to become amused. [He] is probably forgetting about his parents.' [ESᵀ]

---

13 The same meaning can also be coded by the plural form of the count noun A *rand* , ES *tand* 'relative', i.e. A *rand-ɣu*, ES *tand-ɣun* 'relatives'.

(57) ***Xevgun-gu*** *ŋarɣo-t*   *vi-jnə-ɟ.*
     Xevgun-PL   hunt-CVB.NAR.3PL go-DES-IND
     'Xevgun together with other people is going to go hunting.' [A]
     (Panfilov 1962: 113)

2.3.3 Marking of distributivity and iterativity

This section discusses distributivity and iterativity, as marked in nominals. Both categories are discussed in more detail in connection with verbal number (see Section 2.4).

     Noun reduplication can mark the distribution of instrument participants, as ES *paχ~vaχ-kis* 'with stones [one by one]' in (58), or places, as A *eri~eri-in* 'in [various] rivers' in (59). The distribution of other participants is typically marked by the reduplication of the verbal stem. Both in nominal and in verbal reduplicated forms the initial consonant of the second stem alternates essentially according to the general rules of morphophonological alternations.

(58) *azmceɣlŋ e-rχ*   ***paχ~vaχ-kis***   *paz-d.*
     boy   3SG-DAT stone~RED-INSTR throw-IND
     'A boy throws the stones at him [one by one]'. [ES[T]]

(59) *tolvət-ŋan*   *qʰotr* ***eri~eri-in***   *cʰo + ŋəŋ-r*
     be.summer-CVB.SIM bear river~RED-LOC fish + look.for-CVB.NAR.3SG
     i-ɲ-ra ...
     3SG-eat-COORD.3SG
     'In summer, looking for fish in [different] rivers, a bear eats ...' [A]
     (Panfilov 1962: 97)

In the case of nouns with temporal semantics, plain reduplication of noun stems or reduplication in combination with the clitic *=ŋara* is used for deriving iterative adverbs that denote regularly repeated intervals of time, cf. A/ES *muɣf* 'day', *muɣv~muɣf* 'every day', A/ES *parf* 'evening', *parf~parf* 'every evening', A/ES *aɲ* 'year', *aɲ~aɲ=ŋara* 'every year', cf. (60), A/ES *loŋ* 'month', *loŋ~loŋ=ŋara* 'every month'.

(60) ***aɲ~aɲ=ŋara***   *nanχ*   *eɣlŋ + badu-řa-d.*
     year~RED =ITER elder.sister child + give.birth-ITER.3SG.NFUT-IND
     'The elder sister gives/gave birth to a baby every year.' [ES]

In case a noun form comprises the reflexive marker A/ES *pʰ-*, the distributivity is marked by reduplication of the whole stem including the marker, cf. A *pʰ-naχ~pʰ-naχ* 'one's bunks' in (61):

(61) *i-ɲ-tot*   *erk*   *oz-ta*
3SG-eat-CVB.DIST.3PL already get.up-COORD.3PL
pʰ-naχ~pʰ-naχ + *tʰiv-ta.*
REFL-bunk~RED + sit-COORD.3PL
'Having eaten, [they] already stood up, sat down on their bunks...' [= 'each one sat down on his/her own bunk']. [A]
(Nedjalkov and Otaina 2013: 50)

This rule does not, however, apply to the lexical reciprocal marker A/ES *pʰ-ɲafq~ɲafq* <REFL-friend~friend> 'each other', which is derived by reduplicating the root *ɲafq* 'friend' without the reflexive prefix being reduplicated.

    There are also some rare examples like A *ɲivɣ~ɲivɣ-gu* 'people' in (62), where both reduplication and suffixation are used together as markers of plurality. This sentence presumably has a distributive interpretation as well.

(62) *if*   ***ɲivɣ~ɲivɣ-gu*** + *kʰez-ɟ.*
3SG man~RED-PL + tell-IND
'He told [various] people.' [A]
(Panfilov 1962: 97)

2.3.4 Number marking in the vocative

In the Amur variety, Nivkh has two forms of vocative differentiated by number. These forms are typically used in imperative sentences (see Section 3.3), when the speaker directly appeals to the interlocutor. In case of a singular addressee, the vocative forms take the suffix *-a*, cf. *əkən-a* 'elder brother' in (63):

(63) ***əkən-a***   *eɣ-gur*   *pʰrə-ja.*
elder.brother-VOC.SG quickly-ADV.2SG come-IMP.2SG
'Elder brother, come [back] quickly!' [A]
(Saveljeva 1954: 249)

In the plural, the vocative form takes the suffix A *-Ko*, cf. *ɲivɣ-go* 'men' in (64) and *əlc-ɣo* 'slaves' in (65). The plural vocative suffix is transparently of the same origin as the regular plural suffix A *-Ku* and has the same allomorphic distribution (see Section 2.3.1).

(64) ***ɲivɣ-go***   *eɣ-gut*   *mɣe-ve.*
man-VOC.PL quickly-ADV.2PL row-IMP.2PL
'Men, row quickly!' [A]
(Saveljeva 1954: 250)

(65) **əlc-ɣo**   *qʰotr + ŋalivu-da.*
slave-VOC.PL bear + tease-HORT.1PL
'Slaves, let us tease the bear.' [A]
(Panfilov 1965: 223)

In the East Sakhalin variety, the vocative suffix *-a* is used in both numbers. In the singular, it is attached directly to the noun stem in the same way as in the Amur variety, whereas in the plural it comes after the plural suffix, cf. (17) and *andχ-kun-a:* 'guests' in (66):

(66) **andχ-kun-a:** *taf-toχ   j-uɣ-ve.*
guest-PL-VOC house-DAT 3SG-get.in-IMP.2PL
'Guests, come into the house!' [ES]

2.3.5 Number marking in comitative forms

Number plays an essential role in the formation and functioning of the comitative construction that encodes the relationship between participants performing an action together. In the nominal form, a comitative suffix occupies the same position as the plural marker and is combinable with case suffixes. Comitative is expressed cumulatively with number, so Nivkh has the singular and plural comitative forms, whose use is determined by the reference of the noun to which it is attached. The regular plural suffix and both comitative suffixes are obviously historically connected and have the same allomorphic distribution (see Section 2.3.1). Constructions with comitative subjects usually require obligatory marking of plural number on the verbal predicate irrespective of comitative number. Personal pronouns freely take the comitative suffixes according to their own number value (see Section 2.2).

The singular comitative forms in A/L/WS *-Ke*, NS/EST/ESN *-Kin*, EST *-Ken* refer to single participants acting jointly. Both singular and plural comitative suffixes have the same allomorphic distribution as the plural suffix (see Section 2.3.1). The corresponding comitative construction is typically formed by two singular comitative forms, cf. (67)–(69):

(67) *ɲřak tol + mi-ux*   **ius-ke**   **avsq-ɣe**   *va-ɟ-ɣu.*
once water + inside-ABL pike-COM.SG ruff-COM.SG fight-IND-PL
'Once a pike and a ruff fought in the water.' [A]
(Puhta and Otaina 1991: 123)

(68) **əz-gin**   **veχř-kin**   *ɲe-tot*   *um-d-ɣun.*
master-COM.SG servant-COM.SG swear-CVB.DIST.3PL be.angry-IND-PL
'After swearing [at one another] a master and a servant are angry.' [EST]

(69) *miřn* + **ŋafq-ɣen**   ***cʰxəf-ken***   *va-d-ɣun.*
1PL.INCL + friend-COM.SG bear-COM.SG fight-IND-PL
'Our friend and [the] bear fought.' [ES[T]]

The plural comitative forms in A/L/WS -*Ko(n)*, NS/ES -*Kunu* refer to plural partici-pants. In this case, the comitative construction may include either two, cf. (70), or more, cf. (71), plural comitative forms.

(70) **kʰevoŋ-gunu**   **avoŋ-gunu**
villager.upriver-COM.PL villager.downriver-COM.PL
řaŋɢ + *raχc-ɣis*   *v-osqaw-d-ɣun.*
woman + because.of-INSTR REC-fall.out-IND-PL
'The inhabitants of the upriver settlement and the inhabitants of the
downriver settlement fell out over a woman.' [ES[T]]

(71) *harot* *ɲo-rχ*   *təvɣ-r*   **eɲ-yo**   **kʰəs-ko**
then   store-DAT come.in-CVB.NAR.3SG ski-COM.PL ski.stick-COM.PL
puɲ-yo   **kʰu-yo** + *bo-ra*   *rək-r ...*
bow-COM.PL arrow-COM.PL + keep-COORD.3SG bring.in-CVB.NAR.3SG
'Then [he] coming into the store took skis, ski sticks, bows [and] arrows,
bringing [them] in ...' [A]
(Kreinovich 1979: 303)

Mass, singulative-collective and collective nouns (see Section 2.2) always attach the plural comitative suffix. This is illustrated by the plural comitative forms of the mass nouns *raq-ɣunu* 'groats' and *ofa-ɣunu* 'flour' in (72) and *cʰarb-ɣo* in (73), the singulative-collective noun *ŋamx-kunu* in (67), the singulative-collective (paired) nouns *eɲ-yo* 'skis' and *kʰəs-ko* 'ski sticks' in (71), and the collective noun *haɣas-ko* 'clothes' in (73).

(72) *ɲi*   *tuŋ* + *ɲo* + *mi-x*   **raq-ɣunu**   **ofa-ɣunu**   *les*   *si-d.*
1SG this + store + inside-ABL groat-COM.PL flour-COM.PL many put-IND
'I put a lot of groats and flour into the store.' [ES[T]]

(73) *pasq* + *erq* **haɣas-ko**   **cʰarb-ɣo** + *cʰar-ra.*
half + side clothes-COM.PL silk-COM.PL + be.full-COORD.3SG
'Another side is full of clothes and silk.' [A]
(Panfilov 1962: 110)

(74) *qavc* + **ŋamx-kunu**   *urla* + **ŋamx-kunu**   *v-atnaw-d.*
be.gray + hair-COM.PL be.good + hair-COM.PL REC-be.equal-IND
'[The amount of] grey and good hair is equal.' [ES[T]]

In case the referents are pragmatically equal and their number (singular vs. plural) is the same, the comitative suffix is attached to each of the conjoining nominals, cf. (67)–(74). However, in case one or both of these conditions are violated, only the second nominal takes the comitative marker. This happens, first, when the pragmatic balance between the participants is broken, in which case the nominal referring to the more prominent participant comes first and is not marked for comitative. Such a participant is often represented by a personal pronoun, cf. (76)–(77). In such cases the verbal predicate is not marked for plural, cf. (75)–(77):

(75) *ɲ-ŋafq* **pʰ-umgu-ge** *u-ŋətə-ɟ.*
 1SG-friend REFL-woman-COM.SG REC-be.of.the.same.age-IND
 'My friend and his wife are of the same age.' [A]
 (Otaina 1978: 34)

(76) *ɲi* **pʰ-eɣlŋ-gin** *vi-j-d.*
 1SG REFL-child-COM.SG go-FUT-IND
 'I shall go with my child.' [ESᴺ]

(77) *ɲi* **jaŋ-gen** *var-d.*
 1SG 3SG-COM.SG compete-IND
 'I compete with him on sleds.' [ESᵀ]

Second, the corresponding comitative marker is attached only to the second nominal if the number of referents denoted by the nominals is different. Here, two cases are possible. If the first noun has a singular reference, such as *ɲ-əmk* 'my mother' in (78), it is marked neither for number nor for comitative, whereas the second noun with a plural reference, cf. *nanx-kunu* '[and her] elder sisters' in the same example, takes the plural comitative suffix:

(78) *ɲ-əmk* **nanx-kunu** *cʰozjo-t*
 1SG-mother elder.sister-COM.PL whisper-CVB.NAR.3PL
 nudvařklu + aɣraw-d-ɣun.
 something + talk-IND-PL
 'My mother and [her] elder sisters whispering are talking [about] something.' [ESᵀ]

In case the first nominal has a plural reference, such as *ɲiɣvŋ-kun* 'men' in (79), it takes the regular plural suffix, whereas the second nominal with a singular reference, such as *pʰ-aki-xin* '[together with] their elder brother' in the same example, attaches the singular comitative suffix:

(79) *məz* + **ɲiɣvŋ-kun** … **pʰ-aki-xin** *cʰo* + *ŋaŋɣ-t*
2PL.INCL + man-PL   REFL-elder.brother-COM.SG   fish + look.for-CVB.NAR.3PL
vi-χar-t *u-grə-t.*
go-COMPL-CVB.NAR.3PL   REC-be.with.sb-CVB.NAR.3PL
'Our men … together with their elder brother went fishing.' [ES]
(Shternberg 1908: 13)

Comitative nominals may function as a subject, cf. (67)–(70), (74)–(79), a primary object, cf. (71)–(73), an oblique object, cf. (80), or an attribute, cf. (81).

(80) *eɣlŋ-gun paɲřak* **pʰ-ətk-ɣen** **pʰ-əmk-ɣen-doχ**
child-PL often   REFL-father-COM.SG REFL-mother-COM.SG-DAT
pʰřə~pʰřə-d-ɣun.
come~RED-IND-PL
'Children often came to their father and mother.' [ESᵀ]

(81) *ɲi ɲaqr-ux* **kʰeq-xe / kʰeq-xo**
1SG snow-ABL fox-COM.SG / fox-COM.PL
həjk-xe / həjk-xo + *zif-ku* + *ɲřə-ɟ.*
hare-COM.SG / hare-COM.PL + track-PL + see-IND
'I saw fox's / foxes' and hare's / hares' tracks on the snow.' [A]
(Panfilov 1962: 96)

With kinship terms, plural comitative forms can acquire an associative meaning in the same way as regular plural nominal forms (see Section 2.3.2):

(82) **pʰ-acik-xon** **pʰ-nanak-xon**
REFL-younger.sibling-COM.PL REFL-elder.sister-COM.PL
hə + *vo-in* *hum-ɟ-ɣu.*
that:CLOSE + village-LOC be/live-IND-PL
'The younger brother (with his wife) and his elder sister (with her husband) lived in this village.' [A]
(Nedjalkov and Otaina 2013: 58)

2.3.6 Numerals and numeral classifiers

Nivkh has a large class of cardinal numerals, in which it is possible to indicate as high a number as required, at least up to a million. A cardinal numeral may contain either one basic constituent – a numeral element, or two basic constituents – a numeral element and a classifier. The numeral system is decimal, and the enumera-

tion of particular objects is often associated with the semantic features of the object, referred to by classifiers.

Synchronically, classifiers in Nivkh are bound elements and occur only as a part of numerals. There is usually one and only one classifier for a particular noun. Most classifiers are rather transparently derived from noun roots, which means that originally numerals with classifiers constituted a noun phrase. This phrase comprised a modifier represented by a numeral element and a head noun, which was afterwards reanalyzed as a bound classifier. Nivkh used to have at least 33 numeral classifiers that occurred only in the quantification context.[14] Today, most of these classifiers are out of use.

Numerals from 'one' through 'five' obligatorily occur with classifiers, which is why the corresponding numeral elements cannot be used on their own: A/ES *ɲi-ř* <one-CL.SLEDGES> 'one [referring to sledges]', A/ES *mi-m* <two-CL.BOATS> 'two [referring to boats]', A/ES *ca-vr* <three-CL.PLACES> 'three [referring to places]', A/ES *nə-x* <four-CL.ONE.DIM.OBJ.> 'four [referring to one-dimensional objects]', A/ES *tʰo-raχ* <five-CL.TWO.DIM.OBJ.> 'five [referring to two-dimensional objects]', etc.

Numerals referring to the tens normally also take a classifier: A/ES *mχo-vr* <ten-CL.PLACES> 'ten [referring to places], A/ES *cχo-s* <thirty-CL.FISH> 'thirty [referring to fish]'. There are, however, examples in which the corresponding numerals occur without a classifier: *mχo mu* <ten boat> 'ten boats'. Basic numerals over 'five' occur with the classifiers either optionally, cf. A/ES *minr-raχ* <eight-CL.TWO.DIM.OBJ.> 'eight [referring to two dimensional objects]', or never take them, cf. *ŋax řu* <six sledge> 'six sledges.' Classifiers are not attested in the numerals referring to the powers of ten, i.e., 'hundred', 'thousand', and 'million'.

From a semantic point of view, Nivkh classifiers fall into two basic classes. The first class is represented by sortal classifiers, which subsume a counted object under a certain class on the basis of its inherent properties. Sortal classifiers may be ranged according to their specificity/genericity and divided into three subclasses: (a) specific classifiers, which are used in the enumeration of specific objects: ES/A *mu mi-m* <boat two-CL.BOATS> 'two boats'; (b) categoric classifiers, which refer to a category of different objects united on the basis of similar properties: ES *tʰusk ɲe-x* <spruce one-CL.ONE.DIM.OBJ> 'one spruce' (spruce is perceived as a one-dimensional object); and (c) a generic classifier, which refers to the objects that do not enter into any other class: ES *vo nə-kř* <settlement four-CL.GEN> 'four settlements'. The range of semantic oppositions employed in categoric sortal classifiers involves animacy (animates vs. inanimates), being human (humans vs. non-humans), belonging to the class of fish (fish vs. non-fish), ability to form a pair (paired vs. non-paired objects), and extension in the three dimensions (one-dimensional (and small) vs. two-dimensional vs. three-dimensional objects).

---

14 Nivkh numerals and numeral classifiers are discussed in detail in Kreinovich (1932), Panfilov (1953, 1962), Gruzdeva (2004).

Tab. 5: Reconstructed forms and allomorphs of numeral elements.

'ONE'	'TWO'	'THREE'	'FOUR'	'FIVE'
*ɲi	*mi	*ce	*nə	*tʰo
ɲ		c	n	
ɲi	mi			
ɲə			nə	
			nu	
ɲe	me	ce		
ɲa	ma	ca		
	mo			tʰo

The second class is formed by mensural classifiers, which in turn distinguish two subclasses: (a) quantitative classifiers, which establish a unit of measure to indicate quantity of a given object: ES/A *kitə me-fat* <hook two-CL.CORDS.OF.HOOKS> 'two cords of hooks', and (b) extensional classifiers, which are used in measuring extension of an object and are principally based on the old Russian units of measure: ES *tʰom ce-ox* <fat two-CL.FINGERS> 'two fingers of fat'.

A classifier is closely combined or even fused with a numeral element, which sometimes makes them hardly separable. In the process of numeral derivation, both a numeral element and a classifier have undergone various phonological changes. Each of the numeral elements have several attested allomorphs that form correlative rows according to their phonological shape, see Table 5. The choice of a particular allomorph depends on the phonological shape of a classifier following the numeral element. Thus, a final vowel of numeral elements is either dropped or assimilated with vowel(s) of classifiers according to the rules of vowel harmony.

The numeral classifiers, their combinations with the different numerals, their semantic types, and the items being referred to, are summarized in Table 6, with symbols for occurrence (+), non-occurrence (−), or unclear status with regard to occurrence (?) with specific ranges of numerals.

Tab. 6: Nivkh numeral classifiers (based on Gruzdeva 2004).

	CLASSIFIER	SEMANTIC TYPE	ITEMS BEING REFERRED TO	'1'–'5'	'6'–'9'	'10'
1	-ř	specific sortal	sledges	+	–	–
2	-m	specific sortal	boats	+	–	–
3	-u/-i	specific sortal	fishnet cells	+	–	+
4	-řqi/-řqe/-řqəj	specific sortal	fishnet strips	+	?	?
5	-vor/-vur/-fоr	specific sortal	fishnets and fish-spears	+	+	+
6	-o / -u	specific sortal	fishnets for fishing hunchback and Siberian salmon	+	–	+
7	-la/-la/-l	specific sortal	poles for fish-spears	+	+	+
8	-sk	specific sortal	poles for drying fish	+	?	?
9	-c/-ec/-s	specific sortal	big Siberian pipe or spruce boards	+	+	+
10	-řŋ/-řŋk/-ŋiřn, -cʰu/-su	specific sortal	families	+	+	+
11	-svax	specific sortal	generations	+	?	?
12	-vr/-avř	specific sortal	places	+	+	+
13	-x	specific sortal	day's rests on one's way	+	?	?
14	-nŋ/-n, -qř/-qr, -rŋ/-r	categoric sortal, animates, humans	people, good ghosts, gods	+	–	–
15	-n/-ɲ, -ř/-r, -qř/-qr	categoric sortal, animates, non-humans	animals, fish, birds, insects, reptiles, evil ghosts, etc.	+	–	–
16	-s/-os	categoric sortal, animates, fishes	fish	–	–	+
17	-ŋ	categoric sortal, animates, non-fishes	animals, birds, insects, reptiles, evil ghosts, etc.	–	–	+
18	-vasq/-vasq/-vsq/-fasq	categoric sortal, inanimates, pair objects	eyes, ears, hands, legs, sides, skis, oars, mittens, sleeves, etc.	+	+	+

#	Classifier	Category	Objects				
19	-x/-ox	categoric sortal, inanimates, non-pair, one-dimensional objects	sticks, poles, trees, bushes, roots, plants, herbs, roads, brooks, belts, needles, etc.	+	+	−	+
20	-raχ/-raχ/-rax /-rax	categoric sortal, inanimates, non-pair, two-dimensional objects	leaves of plants, pieces of bark, sheets of paper, blankets, sails, robes, shirts, etc.	+	+	+	+
21	-k/-x/-ox	categoric sortal, inanimates, non-pair, three-dimensional objects	bullets, teeth, coins, nuts, beads, eggs, stars, coins, buttons, berries, grains of roe, etc.	+	+	−	+
22	-qř/-oqř/-qr/-oqr/-kř	generic sortal	all other objects	+	+	−	+
23	-qos/-вos/-ɣas	quantitative mensural	special twigs with smelt strung on them	+	+	+	+
24	-ŋaq/-ŋəq	quantitative mensural	twigs with smelt strung on them	+	+	+	+
25	-r/-ar/-arc	quantitative mensural	bundles of slices of dried salmon	+	+	+	+
26	-ŋaq/-rŋaq	quantitative mensural	bundles of dried smelt	+	+	?	?
27	-ɣuvi/-ɣvi/ -xuvi/-χovi	quantitative mensural	bundles of dog's meal	+	+	+	+
28	-rvs	quantitative mensural	bundles of dried grass	+	+	?	?
29	-fat/-fət	quantitative mensural	cords of hooks	+	+	+	+
30	-ux/-ox/-в, -j, -(j)ic	extensional mensural	fingers (as a measure of thickness)	+	+	?	?
31	-ma/-mə	extensional mensural	quarters	+	+	+	+
32	-a/-əa	extensional mensural	sazhens (Rus. *сажень*, a measure of length under 2 m)	+	+	+	+
33	-laj/-laj	extensional mensural	strands of cord	+	+	+	+

Numerals function as the source for distributive adverbs that are derived by reduplicating the numeral, typically 'one', and (optionally) by adding the adverbalizing suffix A -*gur*ː-*gut*, ES -*guř*ː-*gut*. This suffix comprises the causative suffix -*gu*- and the narrative converbal suffix A -*r*ː-*t*, ES -*ř*ː-*t* (see Section 3.3). Examples of the use of distributive adverbs are given in (83)–(85). Note that the reduplicated verbal stem also implies distributivity (see Section 2.4).

(83) *huŋ* + *macala*   *huŋ* + *ɢanŋ*   **ɲa-ɲ~ɲa-ɲ**
    that:CLOSE + guy that:CLOSE + dog one-CL.NON.HUM~RED
    huŋɲi~huŋɲi-d.
    leave.there~RED-IND
    'That guy leaves/left those dogs there one by one.' [ES]

(84) *ɲin*   **ɲe-nŋ~ɲe-nŋ-gut**   *pʰřə~frə-d-ɣun.*
    1PL.EXCL one-CL.HUM~RED-ADV.1PL come~RED-IND-PL
    'We come/came one after another.' [ES]

(85) *cʰχař* + *comř-kun* **ɲa-qr~ɲa-qr-gut**   *kuku-d-ɣun.*
    tree + leaf-PL   one-CL.GEN~RED-ADV.3PL fall.down-IND-PL
    'The tree leaves are falling down one after another.' [ES]

Morphologically and syntactically, numerals are close to nouns. However, unlike nouns they do not inflect for number. The structure of the numeral phrase is discussed in Section 3.2.

2.4 Verbal number

There is a wide range of means expressing the plurality or singularity of events (multiplicative, semelfactive, iterative, distributive) on the verb. These categories are rendered either lexically or morphologically, making use of reduplication and suffixation.[15] There are no traces of suppletion related to verbal number.

Multiplicative plurality, which refers to repeated actions that are performed within the same period of time by the same participant(s),[16] is expressed in Nivkh predominantly by the reduplication of the verbal stem. Here, two derivational paths are possible.

First, multiplicatives can be derived from a simple verbal stem with a semelfactive (momentary) meaning, cf. A/ES *za*- 'knock' > *za~za*- 'knock [repeatedly]', AS/E

---

15 For a more detailed discussion, see Gruzdeva (1997b).
16 The definitions used here for multiplicatives, iteratives and distributives are based on the definitions given in Xrakovskij (1997).

tan- 'nod' > *tan~tan-* 'nod [repeatedly]', A/ES *xici-* 'raise' > *xici~xici-* 'wave', A *varqo-*,
ES *vařku-* 'put on the side' > A *varqo~varqo-*, ES *vařku~vařku-* 'swing [on the waves]',
cf. (86):

(86) *ɲ-aki*   *pʰaχ* + **ca~za-d.**
    1SG-elder.brother window + knock~RED-IND
    'My elder brother knocked [repeatedly] at the window.' [ES]

In case a verb contains the object prefix, the latter is reduplicated together with the
stem: A/ES *j-ama-* <3SG-look> 'look [at sb./sth.]' > *j-am~j-ama-* <3SG-look~3SG-look>
'look [at sb./sth. repeatedly]'. In some rare cases the whole indicative verb form is
reduplicated: A *nə-ɟ* <do-IND> 'do [sth.]' > *nə-ɟ~nə-ɟ* <do-IND~do-IND> 'do [sth. several
times]'.

    Second, multiplicatives can be derived from ideophones whose meaning is con-
nected with visual, auditory and/or tactile perception. The derivation occurs in one
of the two following ways. Two segments of an ideophone can simply be merged
into a single stem and converged into a verb without any additional morphological
marking (but sometimes adding a stem-final vowel), cf. A/ES *tʰuŋ tʰuŋ* 'a sound of
knocking on the wooden object' > *tʰuŋ~tʰuŋa-* 'knock on the wooden object', A/ES
kʰaf kʰaf 'a sound of clanging' > A/ES *kʰaf~kʰaf-* 'clang', A/ES *ŋaw ŋaw* 'mew mew'
> *ŋaw~ŋaw-* 'mew', cf. (87) and more examples in Panfilov (1965: 38).

(87) *jaŋ ŋaɣzəř-kis*   **kʰaf~kʰav-d.**
    3SG tooth-INSTR clang~RED-IND
    'He is clanging his teeth.' [ESᵀ]

Another way to derive a multiplicative verb from an ideophone is to attach the suffix
A *-ju-*, ES *-jo-/-ju-* to the non-reduplicated root of the ideophone, cf. ES *paχ paχ* 'a
sound of clapping' > *paʁ-ju-* 'clap', A/ES *kʰeɣr kʰeɣr* 'a sound of squeaking' > A
kʰeɣr-ju-, ES *kʰeɣr-jo-* 'squeak', A/ES *ŋədr ŋədr* 'a sound of snoring' > A *ŋədr-ju-*, ES
ŋədr-jo- 'snore', cf. (88):

(88) *jaŋ ems-kiř*   **kʰeɣr-jo-d.**
    3SG cheek.tooth-INSTR squeak-V-IND
    'He is gritting his teeth.' [ESᴺ]

Verbs with a semelfactive meaning are derived from lexical multiplicatives by means
of the suffix A *-ɣət-*, ES *-ʁar-*, which otherwise marks a completive or intensive mean-
ing in both varieties and also a resultative meaning in the East Sakhalin variety,[17] cf.

---

17 This suffix is glossed as COMPL.

A *tʰəw-* 'yawn' > *tʰəw-ɣət-* 'yawn [once]', A *jəmr-* 'hiccup' > *jəmr-ɣət-* 'hiccup [once]', ES *ʁov-* 'scoop up' > *ʁov-ʁar-* 'scoop up [once]', ES *odr* 'squeak' > *odr-ʁar-* 'squeak [once]', cf (89):

(89) * řə*   **odr-ʁar-d**.
     door squeak-COMPL-IND
     'The door squeaked [once]. [ES]

In the East Sakhalin variety, semelfactives can also be derived from ideophones by the use of the auxiliary verb *ha-* 'be/do so'[18] which combines with one of the ideophone's segments, cf. *ŋaw ŋaw* 'mew mew' > *ŋaw ha* 'mew [once]', *qʰadr qʰadr* 'a sound of scratching' > *qʰadr ha-* 'scratch [once]', *kʰaf kʰaf* 'a sound of clanging' > *kʰaf ha-* 'clang [once]', cf. (90):

(90) *jaŋ ɲřak=vəřk ŋaɣzəř-kis*   **kʰaf ha-d**.
     3SG once=only tooth-INSTR clang do.so-IND
     'He clanged [his] teeth only once.' [ES^T]

Iterative (habitual) plurality, which refers to repeated actions that are performed at different periods of time by the same participant(s), is marked by the suffix A *-řa-꞉-tʰa-*, ES *-řa-꞉-tʰa-꞉-na-*. From the diachronic point of view, this suffix represents a contamination of the converbal suffix A *-r꞉-t,* ES *-ř꞉-t꞉-n* and the auxiliary verb *ha-* 'be/do so'. This explains the devoicing of *r* > *ř*, which is unusual to the Amur variety, and the aspiration of the allomorph *t* > *tʰ*, which is not typical of the corresponding converbal form. Otherwise, the variants of the iterative suffix are chosen according to the rules of verbal agreement with the subject discussed in Section 3.3, cf. (91) and (92):

(91) *tamk ɲřaklu*   **qo-jo-řa-d**.
     hand sometimes hurt-DIM-ITER.3SG.NFUT-IND
     'The hand sometimes hurts a little bit.' [ES]

(92) *əkəka*   *kə-ba*   *ɲ-əmk + ŋacx-kun*   **qo-tʰa-d-ɣun**.
     every.time rain-CVB.IMMED 1SG-mother + leg-PL hurt-ITER.3PL.NFUT-IND-PL
     'My mother's legs hurt whenever it rains.' [ES]

A sentence with an iterative verb often includes iterative adverbs denoting regularly repeated intervals of time, many of which are formed by the reduplication of the initial noun stem and the clitic *=ŋara* (see Section 2.3.3).

---

18 The verb *ha-* is glossed as 'do so'.

Distributive plurality refers to repeated actions that are performed at the same period of time by non-identical participants. The distribution of participants functioning as subjects or objects is designated in Nivkh by the reduplication of a verbal stem. The distribution of other participants is marked by the reduplication of a nominal stem (see Section 2.3.3). Since the reduplication of the verbal stem alone does not always guarantee the distributive reading, the latter may be supported by the use of the distributive adverbs derived from numerals (see Section 2.3.6). The possibility of using the corresponding distributive adverb in the sentence is actually an indicator of the distributive interpretation of the predicate.

In the subject distributive construction, each of the participants expressed by the subject performs the action successively. The plurality of participants is marked by the plural form of the subject (if the corresponding noun is countable), cf. *ɲiɣvŋ-gun* 'people' in (93), and the plural form of the verbal predicate; whereas their distribution is designated by the reduplication of the verbal stem, cf. *pəkz~pəkz-nd-ɣun* 'disappeared [one after another]':

(93) **ɲiɣvŋ-gun** *palŋ-doχ*   *ŋa + ŋanɣ-t*   *vi-ŋa*
     man-PL   forest-DAT beast + look.for-CVB.NAR.3PL   go-CVB.SIM
     pəkz~pəkz-nd-ɣun.
     disappear~RED-IND-PL
     'When people went hunting into the forest, they disappeared [one after another].' [EST]

In case there is a reason to emphasize that the action has been performed by all participants, the subject is followed by the universal quantifier A *səkm,* ES *sikm* 'all (about humans)', A *sək,* ES *sik* 'all (about non-humans)', and the verbal predicate takes the suffix A *-ɣət-*, ES *-ʁar-*. The meaning of exhaustivity is especially obvious in sentences with a distributive meaning. Thus, in (94), the subject *keχ-kun* 'sea-gulls' bears the plural suffix, the stem of the verb describing the distributive action is reduplicated, and the finite verb takes both of the relevant suffixes: the completive and the plural. The distributive meaning is further reinforced by the adverb *ɲa-n~ɲa-n* 'one by one' and by the use of the quantifier *sik* 'all':

(94) **keχ-kun**   *sik*   *ɲa-n~ɲa-n*   *puj~vuj-t*
     sea.gull-PL   all   one-CL.NON.HUM~RED   fly~RED-CVB.NAR.3PL
     pʰr̆ə-ʁar-d-ɣun.
     come-COMPL-IND-PL
     'The sea-gulls came flying one after another.' [ES]

In the object distributive construction, actions are performed successively with each of multiple participants, cf. (95)–(96), or an action is distributed over various parts of a single participant, cf. (97). In such constructions, distribution is also marked

by the reduplication of the verbal stem and may be additionally emphasized by a distributive adverb, cf. (96).

(95) *qaŋ* *eɣlŋ-gun* + ***duk~ruk-ʁar-d***.
dog child-PL + bite~RED-COMPL-IND
'A dog bit children [all of them, one by one].' [ES]

(96) *əmk* *ŋirŋ-gun* ***ɲa-qr~ɲa-qr-guř*** *zu~zu-d*.
mother cup-PL one-CL.GEN~RED-ADV.3SG wash~RED-IND
'Mother washes the cups one after another.' [ES]

(97) *pil-kař* + *cʰo* *kʰe-roχ* *j-uɣ-ř*
be.big-AUG + fish net-DAT 3SG-get.into-CVB.NAR.3SG
vaχc~vaχc-u-ʁar-d.
tear.up~RED-TRANS-COMPL-IND
'Being caught, the big fish tore up the whole net.' [ES]

An object may also be represented by a mass noun which is used in the singular. In that case, similarly to (97) above, an action is distributed over different portions of a substance:

(98) *tʰom* + ***mařk~mařk-t*** *ma-ta*.
fat + pour.out~RED-CVB.NAR.3PL crush-COORD.3PL
'Pouring out fat [in portions, one by one], [they] crushed [berries].' [ES]

3 Agreement and the syntax of number

This section discusses syntactic and agreement features relevant to number marking, as manifested in a noun phrase (3.1), in a numeral phrase (3.2), and in various types of verbal forms (3.3).

3.1 Number agreement in a noun phrase

Nivkh does not have a category of adjectives, but practically all verbal stems irrespective of their meaning may function as adnominal modifiers that occur in preposition to the head noun. In the Amur variety, an attribute formed from a qualitative verb may express the plurality of the head noun by reduplication, cf. (99). In the East Sakhalin variety, such examples have not been attested.

(99) *imŋ* **pila~pila** + *eri-ux*   *cʰo* + *ŋəŋ-ɟ.*
   3PL be.big~RED + river-ABL fish + look.for-IND
   'They fished in big rivers.' [A]
   (Nedjalkov and Otaina 2013: 50)

Other adnominal modifiers, including demonstratives and possessive pronouns, do not take different forms depending on the number of the head noun.

3.2 Word order in a numeral phrase

Numerals from 'one' through 'five', which obligatorily occur with classifiers (see Section 2.3.6), follow the noun being counted and occupy the postnominal position as is also typical of the universal quantifier A *sək*, ES *sik* 'all (about non-humans)'. In such numeral phrases the noun is normally used in the singular and does not bear any case inflection, which is taken by the numeral, cf. (100). Note that the numeral does not synthesize with the noun being counted and no morphophonological alternations take place at the border of these two words.

(100) *andχ-kun* **mu   me-qř-kiř**   *pʰřə-d-ɣun.*
    guest-PL   boat two-CL.GEN-INSTR come-IND-PL
    'The guests came by two boats.' [ES^N]

All other numerals, including those that optionally occur with the classifiers (see Section 2.3.6), are placed in preposition to the noun being counted. In such phrases the numeral is not declined, whereas the head noun in the singular form bears the case inflection, cf. (101). Again, there is no synthesis between the numeral and the noun being counted.

(101) *andχ-kun* **ŋamg mu-giř**   *pʰřə-d-ɣun.*
    guest-PL   seven boat-INSTR come-IND-PL
    'The guests came by seven boats.' [ES^N]

Panfilov (1962: 98) argues that, probably under influence of Russian, in modern Nivkh there is a new tendency to mark a noun in numeral phrases for plural, cf. *qan-gu* in (102). In my field data, there are also several examples of such pluralization, but I do not consider it to be a strong tendency.

(102) *hə* + *oʁla-gu*   **qan-gu me-qř** + *tʰək-c-ɣu.*
    this:CLOSE + child-PL   dog-PL   two-CL:GEN + bring-IND-PL
    'These children brought two dogs with them.' [A]
    (Panfilov 1962: 100)

3.3 Number agreement of various verb forms

Nivkh verbal predicates agree with the subject in number, conforming to several patterns according to which verbal forms can be divided into three groups. The following section considers the peculiarities of number agreement as manifested in indicative forms (3.3.1), in imperative, hortative and jussive forms (3.3.2), and finally in iterative, coordinated, negative assumptive and apprehensive forms, as well as in some converbs (3.3.3).

3.3.1 Indicative forms

Nivkh indicative forms are derived by the suffix A -ɟ/-c, ES -(n)d/-t, CS -nt and synchronically function as finite predicates. These forms (optionally) agree with the subject in number and in the plural take the same marker as nouns, i.e. A/L/WS -*Ku*, NS/ES/CS/SS -*Kun*. The plural suffix is attached after the indicative marker and may be followed by a discourse clitic: ES *ɲi vi-d=ra* <1SG go-IND=FOC> 'I go/went', *ɲin vi-d-**ɣun**=da* <1PL.EXCL go-IND-PL=FOC> 'we go/went', cf. also example (103) and numerous other examples throughout this paper, especially in Section 2.3. Indicative verb forms are historically nominalizations, which explains the identity of the verbal plural suffix with the nominal one.[19]

The plural marker is fully productive and can be attached to virtually any verbal stem. No agreement in person is attested in the indicative forms:

(103) *cʰχař-kun taf + taʁv-t*   ***pan-d(-ɣun).***
      tree-PL   house + be.around-CVB.NAR.3PL grow-IND-PL
      'Trees grow around the house.' [ES]

Despite the general optional nature of agreement, there are at least two constructions that usually require plural marking on the verbal predicate. One of them is the comitative construction (see Section 2.3.5), while the other is the reciprocal one, cf. (104). There are, however, additional factors (not discussed in the present paper) that may override this obligatory agreement.

(104) *imŋ **u-ŋala-ɟ-ɣu**.*
      3PL   REC-resemble-IND-PL
      'They resemble each other.' [A]
      (Nedjalkov and Otaina 2013: 108)

---

19 Until recently a predicate of a clause was apparently comprised of a nominal form in A -ɟ/-c, ES -(n)d/-(n)t and an auxiliary verb *ha-* 'be / do so' marked by one or another mood suffix. Later, the auxiliary was lost in the epistemically neutral clauses and lexicalized into a modal particle in the

3.3.2 Hortative, imperative and jussive forms

The only finite verb forms in Nivkh that display differentiation of person and number are hortatives, imperatives and jussives.[20] They are formed by attaching the suffixes listed in Table 7 to a verbal stem.

In the first person (= hortatives), three numbers (singular, dual, and plural) are differentiated in the Amur variety and two numbers (dual and plural) in the East Sakhalin variety. The first person singular form is attested only in the Amur variety:

(105) *nama-gut   cəj   kʰrəu-tot   **pʰu-nəkta.***
 be.good-ADV.1SG again rest-CVB.DIST.1SG go.out-HORT.1SG
 'After having a good rest, let me go out!' [A]
 (Panfilov 1965: 131)

In a similar context the East Sakhalin speakers traditionally use the future indicative form: ES *ra-j-d=ra* <drink-FUT-IND=FOC> 'Let me drink / I shall drink'. Interestingly, modern speakers of this variety have started using two new analytical first person forms that are structurally similar to those existing in Russian and are apparently a result of Russian-Nivkh interference (Gruzdeva 2000). These complex forms consist of two components. The first component is represented by the special petrified imperative forms *tʰana* <give:IMP.2SG> and *tʰana-ve* <give:IMP.2SG-IMP.2PL>, which can still be used as regular lexical verbs with the meanings 'give [you:SG]!' and 'give [you:PL]!', respectively. The second component is a regular imperative

Tab. 7: Hortative, imperative and jussive suffixes.

		SINGULAR	DUAL	PLURAL
1 PERSON (HORTATIVE)	A	*-nəkta, -nəxta*	*-nəte, -nte*	*-da*
	ES	–	*-nate*	*-da*
2 PERSON (IMPERATIVE)	A	*-ja, -j*	*-Ve*[21]	
	ES	*-ja, -j*	*-Ve*	
3 PERSON (JUSSIVE)	A	*-ʁazo*		
	ES	*-ʁaro*	*-ʁarʁ(a)ro*	

---

epistemically marked clauses. The suffix A -ɟ/-c, ES -(n)d/-t was reanalyzed as an indicative mood suffix and the forms with this suffix started being used as finite verb forms (Gruzdeva 2016).
20 For more details, see Panfilov (1965), Kreinovich (1979), Gruzdeva (2001), Nedjalkov and Otaina (2013).
21 The imperative suffix *-Ve* has the allomorphs *-ve/-pe/-ɓe*.

form (either singular or plural) with the causative suffix -*gu*-. The resulting forms are opposed according to the number of hearers:

(106) a. ***tʰana***   *ɲ-aχ*   ***lu-gu-ja***.
     give:IMP.2SG 1SG-CAUSEE sing-CAUS-IMP.2SG
     'Let me sing [you:SG]!' [ES]

     b. ***tʰana-ve***   *ɲ-aχ*   ***lu-gu-ve***.
     give:IMP.2SG-IMP.2PL 1SG-CAUSEE sing-CAUS-IMP.2PL
     'Let me sing [you:PL]!' [ES]

The first person dual (inclusive) form, cf. A/WS *ra-nəte*, ES *ra-nate* <drink-HORT. 1DU> 'let us [two] drink' (107), has practically vanished,[22] the first person plural form being used instead.

(107) *pʰ- řuɣř + coz-u-ř*   ***pʰim-nəte***.
     REFL-fire + put.out-TR-CVB.NAR.1DU run.away-HORT.1DU
     'Putting out our fire, let us [you:SG and me] run away!' [WS]
     (Panfilov 1965: 130)

The first person plural (inclusive) hortative form has the same marker -*da* < *-*nta* in all varieties, cf. A/ES *ra-da* <drink-HORT.1PL> 'let us drink' and (108).

(108) *enf-toχ*   *vi-t*   ***hum-da***.
     another.place-DAT go-CVB.NAR.1PL be-HORT.1PL
     'Going to another place, let us live.' [A]
     (Panfilov 1965: 130)

The singular and plural second person imperative markers are also identical in all varieties: A/ES *ra-ja* <drink-IMP.2SG> 'drink [you:SG]' and (109), A/ES *ra-ve* <drink-IMP.2PL> 'drink [you:PL]'and (110):

(109) *nav-ux*   *taf-toχ*   ***vi-ja***.
     now-ABL house-DAT go-IMP.2SG
     'Go to the house now!'[ES]

(110) *hana macala-go*   *parvət-ŋan*   ***pʰrə-ve***.
     well   guy-VOC.PL be.evening-CVB.SIM come-IMP.2PL
     'Well, guys, come in the evening!' [A]
     (Panfilov 1965: 130)

---

22 The same has happened to the corresponding first person dual personal pronoun.

As for the third person jussive forms, in the Amur variety only one form is used both for singular and plural: A *if ra-ʁazo* <3SG drink-JUS.3SG/PL> 'let him/her drink', *imɣ ra-ʁazo* <3PL drink-JUS.3SG/PL> 'let them drink'. Judging by earlier data, the same underdifferentiation of number in the third person used to be observed in the East Sakhalin variety as well. The number opposition in the third person imperative forms, with the reduplicated form of the suffix being used in the plural, cf. (111)–(112), has been attested only in my own data.

(111) *ɲaŋ   nana oz-ba   iŋk + **aj-ʁaro**.*
3SG   now   get.up-CVB.IMMED food + make-JUS.3SG
'Now as soon as she gets up, let her make food.' [ES]

(112) *ɲi   nana pʰɹ̌ə-ba   in   ɲe-rχ   **pʰɹ̌ə-ʁarʁro**.*
1SG now   come-CVB.IMMED 3PL 1SG-dat come-JUS.3PL
'Now as soon as I come, let them come to me.' [ES]

3.3.3 Converbs and other verb forms

Nivkh has a whole set of finite and non-finite (converbal) forms that agree with the subject in person and number, cf. Table 8.[23] All markers of these forms include the same elements *r* and *t* in the Amur variety and *r/ř*, *t*, and *n* in the East Sakhalin

Tab. 8: Markers of verbal forms that agree with the subject in person and number.

	A	ES	
I	-ra : -ta	-(ŋ)ra : -(ŋ)ta	emphatic/evidential mood (finite)
	-rla : -tla	-rlo : -tlo	assumptive mood (finite)
	-nəra : -nəta?	-inəŋra : -inəŋta	visual apprehensive mood (finite)
	-(i)jra : -(i)jta?	-jaŋra : -jaŋta	inferential apprehensive mood (finite)
	-datar : -datat	(-data)	temporal converb (bounded simultaneity)
	-durŋur : -durŋut	–	temporal converb (bounded simultaneity)
	-vur : -vut	-vuř : -vut	quotational converb
	-iləkrr : -iləkrt	-ilakrř : -ilakrt	negative-purposive converb
II	-ř̌a- : -tʰa-	-ř̌a- : -tʰa- : -na-	iterative suffix
	-ra : -ta	-ra : -ta : -na	coordination (finite)
	-r : -t	-ř : -t : -n	narrative converb
	-ror : -tot	-roř : -tot : -non	temporal converb (distant anteriority)
	-gur : -gut	-guř : -gut : -gun	purposive converb

23 This topic was first examined in Kreinovich (1983) and then discussed in more details in Gruzdeva (1998); Nedjalkov and Otaina (2013).

Tab. 9: Distribution of common elements of suffixes (*r* , *ř* vs. *t*) in A and ES.

	SINGULAR	PLURAL
1 PERSON	*t*	*t*
2 PERSON	*r* , *ř*	*t*
3 PERSON	*r* , *ř*	*t*

variety (the voiceless trill *ř* occurs word-finally). The iterative suffix is somewhat exceptional due to its historical origin (see Section 2.5). Verb forms that agree with the subject can be divided into two groups.

Group I in Table 8 comprises eight verb forms, each of which in both varieties has two variants of suffixes – one with the consonant *r* or *ř* and another with the consonant *t*. The alternation of these consonants indicates the agreement of the corresponding verb form with the subject in person and number according to the pattern presented in Table 9. As one can see, the consonant *r* or *ř* is attested in the second and third person singular forms, whereas the consonant *t* appears in all other persons and numbers.

The alternation of consonants can synchronically take place (a) suffix-initially, cf. A/ES -*ra*꞉-*ta*, A -*rla*꞉-*tla*, ES -*rlo*꞉-*tlo*; (b) suffix-medially, cf. A -*nəra*꞉-*nəta*ˀ, ES -*inəŋra*꞉-*inəŋta*, A -*ijra*꞉-*ijta*ˀ, ES -*jaŋra*꞉-*jaŋta*; and (c) suffix-finally, cf. A -*datar*꞉-*datat,* A -*durŋur*꞉-*durŋut,* A -*iləkrr*꞉-*iləkrt,* ES -*ilakrř*꞉-*ilakrt,* A -*vur*꞉-*vut,* ES -*vuř*꞉-*vut*. As will be seen from the further discussion of concrete forms, most of the listed suffixes are compositional and consist of two or more elements.

The emphatic/evidential forms express direct (visual) evidentiality and are marked by the suffix A -*ra*꞉-*ta,* ES -*(ŋ)ra*꞉-*(ŋ)ta*, which occupies the same slot as other mood suffixes. The agreement of these forms with the subject is illustrated in (113)–(114). The emphatic/evidential forms are rarely used in the modern language, but they serve as a basis for deriving several other mood forms.

(113) *ɲ-ontq=aʁr c^h-ontq + voci-ř* ***per-ra.***
1SG-bag=FOC 2SG-bag + be.similar-CVB.NAR.3SG be.heavy-EVID.DIR.3SG
'My bag is as heavy as your bag.'[ES^T]

(114) *naf=aχ meŋ eχlaŋ=aʁr +* ***iv-ifu-ta.***
now=FOC 1DU child=FOC + have-PROGR-EVID.DIR.1DU
'Now we two are going to have a child.' (A bear-woman is telling to a man) [ES]
(Kreinovich 1979: 316)

Assumptive forms are typically used in the future tense and indicate that the situation will apparently not take place, because it usually does not occur under similar circumstances. These forms bring the nuance of uncertainty. The marker is the com-

positional suffix A *-rla : -tla,* ES *-rlo : -tlo,* which is transparently composed of the narrative converbal suffix *-r : -t* and the probabilitive clitic *=lo*:

(115) a. *jaŋ **pʰřə-j-rlo**.*
     3SG come-FUT-ASS.3SG
     'He will probably not come.' [ES]

   b. *ɲin **pʰřə-j-tlo**.*
     1PL.EXCL come-FUT-ASS.1PL
     'We will probably not come.' [ES]

Apprehensive forms denote the speaker's attempt to prevent undesirable consequences of the situation. Visual (direct) apprehensive is issued in the situation when the speaker directly observes the addressee's behavior that may lead to undesirable consequences. In the Amur variety it is marked by the compositional suffix A *-nəra*, cf. (116), which may also have the variant *-nəta*[?]. The suffix consists of the future tense suffix *-nə-* and the emphatic/evidential suffix *-ra : -ta.*

(116) *əlc-ɣo eʁ-gut oz-ve. qʰotr **vi-nəra**.*
servant-VOC.PL be.quick-ADV.2PL get.up-IMP.2PL bear go-APPR.VIS.3SG
'Servants, get up quickly! [Otherwise] a bear will go away!' [A]
(Panfilov 1965: 118)

In the East Sakhalin variety, the corresponding visual apprehensive form is derived with the suffix *-inəŋra : -inəŋta*, which is composed of the desiderative suffix *-inə-* and the emphatic/evidential suffix *-(ŋ)ra : -(ŋ)ta.* For instance, when the hearer(s) is/are going through a slippery road and may fall down, the speaker uses the following forms:

(117) a. ***kuc-inəŋra**.*
     fall.down-APPR.VIS.2SG
     'Do not fall down!' [ES]

   b. ***kuc-inəŋta**.*
     fall.down-APPR.VIS.2PL
     'Do not fall down!' [ES]

Inferential apprehensive is used when the speaker does not directly observe any potentially harmful actions carried out by the hearer but, proceeding from previous negative experience, assumes that in the future such actions may take place. In the Amur variety the marker *-(i)jra* consists of the emphatic/evidential suffix *-ra* and the element *-(i)j-*, which seems to indicate some kind of modality, cf. (118). Again, it is likely that there does exist the variant *-(i)jta*[?] of the same suffix.

(118) *cʰi vəɣdr-r*   *e-v-ja*   *vəj-rox*
       2SG catch.on.the.fly-CVB.NAR.2SG 3SG-keep-IMP.2SG bottom-DAT
       kut-ku-jra.
       fall.down-CAUS-APPR.INF.2SG
       'Catch [it] on the fly, do not let [it] fall down.' [A]
       (Panfilov 1962: 149)

In the East Sakhalin variety the corresponding inferential apprehensive suffix is *-jaŋra* : *-jaŋta*. Both variants comprise the same component *-ja-*, which can probably be connected with the second person singular imperative marker *-ja*, and the emphatic/evidential suffix *-(ŋ)ra* : *-(ŋ)ta*. The forms presented in (119) are used in the situation when the speaker is giving a fragile object to the hearer(s):

(119) a. ***zosq-jaŋra.***
         break-APPR.INF.2SG
         'Don't break [it]!' [ES]

       b. ***zosq-jaŋta.***
         break-APPR.INF.2PL
         'Don't break [it]!' [ES]

The markers of the temporal converbs in A *-datar* : *-datat,* A *-durŋur* : *-durŋut* with more or less similar meaning of bounded simultaneity are also compositional and comprise the elements *-data-* and *-durŋu-* (apparently of verbal origin) and the narrative converbal suffix *-r* : *-t*, which indicates the agreement with the subject:

(120) *kʰeq* ***vi-datar*** / ***vi-durŋur*** *mu-ɟ.*
       fox   go-CVB.BSIM.3SG   die-IND
       'The fox [fell] dead while going.' [A]
       (Nedjalkov and Otaina 2013: 331)

The peculiarity of the quotational converb in A *-vur* : *-vut,* ES *-vuř* : *-vut* is that it agrees not with its "own" subject, but, rather, with the subject of the finite verb, cf. (13) and (121). The agreement follows the pattern in Table 9.

(121) *cʰo + řom* ***ha-vut***   *it-t*   *in-aχ*
       fish + oil   do.so-CVB.QUOT.3PL say-IND 3PL-CAUSEE
       ra-gu-tʰa-d-ɣun.
       drink-CAUS-ITER.3PL.NFUT-IND-PL
       '[He] said [that] they drank them with fish oil.' [ES]

The negative-purposive converb is formed with another compositional marker A *-iləkrr* : *-iləkrt,* ES *-ilakrř* : *-ilakrt,* which comprises the synchronically bound elements A *-ləkr-,* ES *-lakr-* and the narrative converbal suffix A *-r* : *-t*, ES *-ř* : *-t*. This

converb also agrees with the subject of the finite verb, which is illustrated by example (122). In (122a), *ra-jləkrt* 'so that [they] would not drink' agrees with the subject of the finite verb *ɲi* 'I', whereas in (122b), *ra-jləkrr* with the same meaning agrees with the subject of the finite verb *if* 'he':

(122) a. *ɲi   e-f-t*   *řorvi-ɉ=ra*   *o:la-gu*
       1SG   3SG-keep-CVB.NAR.1SG carry.away-IND=FOC child-PL
       ra-jləkrt.
       drink-CVB.NEG.PURP.1SG
       'Taking it, I carried [the medicine] away so that the children would not drink [it].' [A]

     b. *if   e-v-r*   *řorvi-ɉ=ra*   *o:la-gu*
       3SG   3SG-keep-CVB.NAR.3SG carry.away-IND=FOC child-PL
       ra-jləkrr.
       drink-CVB.NEG.PURP.3SG
       'Taking it, he carried [the medicine] away so that the children would not drink [it].' [A]
       (Kreinovich 1979: 321)

In group II (see Table 8), the Amur verb forms have two variants of suffixes with the consonants *r* or *t* (similarly to the forms from group I), whereas the corresponding East Sakhalin forms have three variants with the consonants *r / ř, t* or *n*. In this group, the alternation of consonants takes place (a) suffix-initially, cf. A *-řa-* : *-tʰa-*, ES *-řa-* : *-tʰa-* : *-na-*, A *-ra* : *-ta*, ES *-ra* : *-ta* : *-na*, A *-r* : *-t*, ES *-ř* : *-t* : *-n;* (b) suffix-finally, cf. A *-gur* : *-gut*, ES *-guř* : *-gut* : *-gun*; and (c) both suffix-initially and suffix-finally, cf. A *-ror* : *-tot*, ES *-roř* : *-tot* : *-non*. The Amur forms demonstrate the same agreement pattern, as the forms from group I (see Table 9). In the East Sakhalin variety, the situation is more complicated. Here, the second and third person singular verb forms always comprise the consonant *r / ř*, in the same way as the forms from group I. In all other numbers and persons, the choice between variants with *t* and *n* depends on the mood/aspect of the verb form, in which the suffix occurs, or, in case of converbs, is determined by the mood/aspect of the finite verb. In case the (finite) verb occurs in the future tense, optative, hortative, imperative or jussive mood, the verb forms in question begin and/or end in the consonant *n*, as shown in Table 10.

Tab. 10: Distribution of common elements of suffixes (*r / ř* and *n*) in ES.

	SINGULAR	PLURAL
1 PERSON	*n*	*n*
2 PERSON	*r , ř*	*n*
3 PERSON	*r , ř*	*n*

In other tenses and moods, the corresponding forms begin and/or end in the consonant *t*, as shown in Table 9.

The iterative suffix A *-řa-*꞉*-tʰa-*, ES *-řa-*꞉*-tʰa-*꞉*-na-* syncretically expresses iteration, person and number (in all varieties), as well as tense/mood of the verb form, in which it occurs (in the East Sakhalin variety), cf. (91)–(92), (121), (123).

(123) a. *ɲi*   ***pʰřə-tʰa-d.***
           1SG come-ITER.1SG.NFUT-IND
           'I came/come [repeatedly].' [ES]

    b. *in*   ***pʰřə-tʰa-d-ɣun.***
           3PL come-ITER.3PL.NFUT-IND-PL
           'They came/come [repeatedly].' [ES]

    c. *jaŋ*   ***pʰřə-řa-d.***
           3SG come-ITER.3SG.FUT/NFUT-IND
           'S/he came/come/will come [repeatedly].' [ES]

    d. *ɲi*   ***pʰřə-na-d.***
           1SG come-ITER.1SG.FUT-IND
           'I shall come [repeatedly].' [ES]

    e. *in*   ***pʰřə-na-d-ɣun.***
           3PL come-ITER.3PL.FUT-IND-PL
           'They will come [repeatedly].' [ES]

In the imperative, the variant of the iterative suffix is chosen in the Amur variety according to Table 9, cf. (18), and in the East Sakhalin variety according to Table 10, cf. (124):

(124) a. ***pʰřə-řa-ja.***
           come-ITER.2SG.IMP-IMP.2SG
           'Come [you:SG repeatedly]!' [ES]

    b. ***pʰřə-na-ve.***
           come-ITER.2PL.IMP-IMP.2PL
           'Come [you:PL repeatedly]!' [ES]

Coordinated forms appear as predicates of chaining coordinated clauses. Each coordinated form takes the same set of suffixes A *-ra*꞉*-ta*, ES *-ra*꞉*-ta*꞉*-na* with the variants chosen according to the person and number of the subject in both varieties and the tense/mood of the verb form in the East Sakhalin variety, cf. Tables 9 and 10, examples (16), (26), (61) and (125). The chain of coordinated forms may be closed by the auxiliary verb *ha-* 'be / do so', which is optional in the indicative and obligatory in other moods:

(125) a. *jaŋ* *ə̌rk* **pil-ra** **vesqar-ra** (*ha-d*).
3SG already be.big-COORD.3SG be.strong-COORD.3SG do.so-IND
'He is already big and strong.' [ES]

b. *ɲi* *ə̌rk* **pil-ta** **vesqar-ta** (*ha-d*).
1SG already be.big-COORD.1SG.NFUT be.strong-COORD.1SG.NFUT do.so-IND
'I am already big and strong.' [ES]

c. *ɲi* **pil-na** **vesqar-na** (*ha-j-d*).
1SG be.big-COORD.1SG.FUT be.strong-COORD.1SG.FUT do.so-FUT-IND
'I shall be big and strong.' [ES]

The narrative converb in A *-r* : *-t*, ES *-ř* : *-t* : *-n* and the temporal converb in A *-ror* : *-tot*, ES *-roř* : *-tot* : *-non* agree with the subject of the finite verb. In case the subjects of a converb and a finite verb are non-coreferential, the converb takes the causative suffix *-gu-*, which functions as a switch-reference marker. Otherwise, given the differences between the language varieties, the agreement takes place according to the person and number of the subject and the tense/mood of the finite verb, as represented in Tables 9 and 10:

(126) a. *jaŋ* *cʰo* + **xu-roř** *kʰə* + **ye-ř** *vi-d*.
3SG fish + kill-CVB.DIST.3SG axe + take-CVB.NAR.3SG go-IND
'After catching fish, taking the axe, he went.' [ES]

b. *ɲi* *cʰo* + **xu-tot** *kʰə* + **ye-t** *vi-d*.
1SG fish + kill-CVB.DIST.1SG.NFUT axe + take-CVB.NAR.1SG.NFUT go-IND
'After catching fish, taking the axe, I went.' [ES]

c. *cʰin* *cʰo* + **xu-non** *kʰə* + **ye-n** *vi-ve*.
2PL fish + kill-CVB.DIST.2PL.IMP axe + take-CVB.NAR.2PL.IMP go-IMP.2PL
'After catching fish, taking the axe, go [you:PL]!' [ES]

d. *jaŋ* *cʰo* + **xu-gu-tot** *ɲi* *vi-d*.
3SG fish + kill-CAUS-CVB.DIST.1SG.NFUT 1SG go-IND
'After he had caught fish, I went.' [ES]

When the causative suffix *-gu-* is added to a narrative converbal stem, the resulting form, i.e., A *-gur* : *-gut,* ES *-guř* : *-gut* : *-gun*, often acquires a purposive meaning and functions as a purposive converb with a complex marker:

(127) *tʰolf* **qʰav-qʰatn-guř** *cŋəř* + *mařqa-d*.
summer be.hot-AUG-CVB.PURP.3SG grass + water-IND
'Because summer was very hot, [she] watered the grass.' [ES]

4 Semantics and discourse

This section examines the phenomenon which is connected with such pragmatic features as honorifics and politeness, as indicated by the verbal imperative forms.

In general, Nivkh does not have special honorific forms or registers. The only forms that may have the honorific connotation are the imperative ones – at least this is the case in the East Sakhalin variety.

Appealing to a single elder hearer, the Nivkh speaker may traditionally use an honorific imperative verb form, i.e. the second person plural imperative form in *-Ve* (see Section 3.3.2) containing the additional second plural iterative suffix in the form of *-na* (see Section 3.3.3), cf. *vi-na-ve* <go-ITER.2PL.IMP-IMP.2PL> 'go [you:SG]'. This form, which can also be understood as "an iterative plural", is perceived as a very polite one.

Nowadays, a polite request may also be expressed by the standard second person plural imperative form without the iterative suffix, cf. *vi-ve* <go-IMP.2PL> 'go [you:SG]' and *um-ɢavr-ve* 'do not be angry [you:SG]' with an honorific singular reference in (128). Such a way of expressing politeness has obviously developed under the influence of Russian.

(128) *andχ-a* **um-ɢavr-ve** *ur-gu-n*
  guest-VOC be.angry-NEG-IMP.2PL be.good-CAUS-CVB.NAR.1PL.HORT
  ŋafq-ɣun + voci-n *keraj-da.*
  friend-PL + be.like-CVB.NAR.1PL.HORT speak-HORT.1PL
  'Guest! Do not be angry [you:SG], let us speak as good friends!' [ES]

5 Conclusions

As can be seen from the previous discussion, number is manifested in Nivkh as several different types of phenomena. In general terms it may be considered as (a) an inherent property of pronouns, (b) an inflectional nominal category (marked by suffixation), (c) an inflectional verbal category based on agreement with the subject (marked by suffixation in a set of finite and converbal forms), and (d) a component of a clausal quantitative aspectual structure (marked by verbal suffixation or reduplication and by reduplication of nominal dependents).

A maximum number of number features (singular vs. dual vs. plural) is attested in the first person pronouns and in the hortative verb forms, whereas in all other persons, lexical classes and grammatical forms a two-way number contrast between singular and plural is observed. Basic numerals and numerals referring to tens typically comprise a numeral element and a suffixal classifier, which can be seen as an areal feature with an East and South Asian distribution.

Number is not obligatory as an inflectional category either for nouns or for verbs. Inflectional verb forms display various patterns of verbal agreement. Coordinated forms, some other marginal finite forms, as well as a group of converbs, agree with the subject in person and number in a completely different way than both indicative forms on the one hand and hortative, imperative and jussive forms, on the other. Such a complex picture can best be explained by assuming that the verbal forms are results of morphosyntactic restructuring at different time levels. At the deepest level, Pre-Proto-Nivkh seems to have been a rather analytic language that underwent a process of "Altaicization", gradually shifting towards the structures typical of the neighboring Tungusic languages. As a result of interaction with its Altaic-type neighbors, Nivkh has acquired its overall present morphosyntactic system, agglutinative morphology, the finite use of nominalizations and the wide use of converbs (Gruzdeva and Janhunen 2020).

Acknowledgements

List of Abbreviations

ABL	ablative
ADV	adverb
APPR	apprehensive
ASS	assumptive
AUG	augmentative
BSIM	bounded simultaneity
CAUS	causative
CL	classifier
COM	comitative
COMPL	completive
COORD	coordinated
CVB	converb
DAT	dative
DES	desiderative
DEST	destinative
DIM	diminutive
DIR	direct
DIST	distal
DU	dual
DUR	durative

EVID	evidential
EXCL	exclusive
FOC	focus
FUT	future
HORT	hortative
IMMED	immediate
IMP	imperative
INCL	inclusive
IND	indicative
INSTR	instrumental
ITER	iterative
JUS	jussive
LOC	locative
MOD	modal
NAR	narrative
NEG	negative
NMLZ	nominalizer
NFUT	non-future
PL	plural
PROB	probabilitive
PROGR	progressive
PURP	purposive
Q	question
QUAL	qualitative
QUOT	quotative
REC	reciprocal
REFL	reflexive
SIM	simultaneity
TRANS	transitive
V	verbalizer
VIS	visual
VOC	vocative

References

Austerlitz, Robert. 1959. Semantic components of pronoun systems: Gilyak. *Word* 15 (1). 102–109.

Gruzdeva, Ekaterina. 1997a. Aspects of Nivkh morphophonology: Initial consonant alternation after sonants. *Journal de la Société Finno-Ougrienne* 87. 79–96.

Gruzdeva, Ekaterina. 1997b. Plurality of situations in Nivkh. In Viktor S. Xrakovskij (ed.), *Typology of Iterative Constructions* (LINCOM Studies in Theoretical Linguistics 4), 164–185. München; Newcastle: LINCOM Europa.

Gruzdeva, Ekaterina. 1998. *Nivkh*. (Languages of the World: Materials 111). München; Newcastle: LINCOM Europa.

Gruzdeva, Ekaterina. 2000. Aspects of Russian-Nivkh grammatical interference: the Nivkh imperative. In Dicky Gilbers, John Nerbonne & Jos Schaeken (eds.), *Languages in Contact* (Studies in Slavic and General Linguistics 28), 121–134. Amsterdam – Atlanta, GA: Rodopi.

Gruzdeva, Ekaterina. 2001. Imperative sentences in Nivkh. In Viktor S. Xrakovskij (ed.), *Typology of Imperative Constructions* (LINCOM Studies in Theoretical Linguistics 9), 59–77. München; Newcastle: LINCOM Europa,

Gruzdeva, Ekaterina. 2004. Numeral classifiers in Nivkh. In Alexandra Y. Aikhenvald (ed.), *Language Typology and Universals. Nominal Classification*. *Sprachtypologie und Universalienforschung* 57 (2/3). 300–329.

Gruzdeva, Ekaterina. 2016. Epistemic modality and related categories in Nivkh. In Ekaterina Gruzdeva & Juha Janhunen (eds.), *Crosslinguistics and linguistic crossings in Northeast Asia. Papers on the languages of Sakhalin and adjacent regions*, 171–198. (Studia Orientalia 117). Helsinki.

Gruzdeva, Ekaterina & Juha Janhunen. 2020. Notes on the typological prehistory of Ghilyak. *International Journal of Eurasian Linguistics* 2 (1). 1–28.

Kreinovich, Eruhim A. 1932. *Giljackie chislitel'nye* [Giljak numerals] (Trudy nauchno-issledovatel'skoj associacii Instituta narodov Severa CIK SSSR I.3). Leningrad.

Kreinovich, Eruhim. 1937. Fonetika nivxskogo (giljackogo) jazyka [Nivkh (Giljak) phonetics]. 7–102. Moskva; Leningrad: Institut narodov Severa / Uchpedgiz,

Kreinovich, Eruhim A. 1934. Nivxskij (giljackij) jazyk [The Nivkh (Giljak) Language]. In Eruhim A. Krejnovich (ed.), *Jazyki i pis'mennost' narodov Severa III. Jazyki i pis'mennost' paleoaziatskix narodov,* 181–222. Moskva; Leningrad: Uchpedgiz.

Kreinovich, Eruhim A. 1979. Nivxskij jazyk [The Nivkh language]. In *Jazyki Azii i Afriki III. Jazyki drevnej perednej Azii (nesemitskie). Iberijsko-kavkazskie jazyki. Paleoaziatskie jazyki,* 295–329. Moskva: Nauka.

Kreinovich, Eruhim A. 1983. Ob odnoj unikal'noj konstrukcii slozhnopodchinennogo predlozhenija v nivxskom jazyke [On a unique subordinate sentence construction in Nivkh]. In: Jurij Sem & Alexandr I. Gamilov (eds.), *Voprosy leksiki i grammatiki jazykov narodov Krajnego Severa SSSR*, 105–116. Leningrad: LGPI.

Mattissen, Johanna. 2003. *Dependent-head synthesis in Nivkh: a contribution to a typology of polysynthesis*. Amsterdam/Philadelpia: John Benjamins.

Nedjalkov, Vladimir P. & Galina A. Otaina. 2013. *A Syntax of the Nivkh language. The Amur dialect*. Ed. by Emma Š. Geniušiene & Ekaterina Gruzdeva. Amsterdam, Philadelphia: John Benjamins.

Otaina, Galina A. 1978. *Kachestvennye glagoly v nivxskom jazyke* [Qualitative verbs in Nivkh]. Moskva: Nauka.

Panfilov, Vladimir Z. 1953. *Nivxskie kolichestvennye chislitel'nye* [Nivkh cardinal numerals]. Avtoreferat dissertacii na soiskanie uchjonoj stepeni kandidata filologicheskix nauk. Leningrad.

Panfilov, Vladimir Z. 1962. *Grammatika nivxskogo jazyka* [A grammar of Nivkh] 1. Moskva, Leningrad: Nauka.

Panfilov, Vladimir Z. 1965. *Grammatika nivxskogo jazyka* [A grammar of Nivkh] 2. Moskva, Leningrad: Nauka.

Puhta, Marija N. & Galina A. Otaina. 1997. *Nivxgu dif* (Uchebnik i kniga dlja chtenija dlja 2 klassa (amurskij dialekt)). [Nivkh Word. Textbook for grade 2 (Amur dialect). Leningrad.

Sangi, Vladimir M. & Ljudmila B. Gashilova. 2003. *Nivxsko-russkij slovar'* [Nivkh-Russian dictionary]. Sankt-Peterburg: Prosveschenie.

Saveljeva, Valentina N. 1954. Sklonenie imen suschestvitel'nyx v nivxskom jazyke. [Nivkh noun declension]. *Uchenye zapiski LGPI* 101. 239–254. Leningrad: LGPI.

Saveljeva, Valentina N. & Chuner M. Taksami. 1970. *Nivxsko-russkij slovar'* [Nivkh-Russian dictionary]. Moskva: Sovetskaja enciklopedija.

Shternberg, Lev Ja. 1908. *Materialy po izucheniju giljackogo jazyka i fol'klora I. Obrazcy narodnoj slovesnosti I* [Materials for the study of the Giljak language and folklore I. Samples of oral literature I]. (Izvestija imperatorskoj Akademii nauk XXII). Sankt-Peterburg.

Xrakovskij, Viktor S. (ed.) 1997. *Typology of iterative constructions* (LINCOM Studies in Theoretical Linguistics 4). München, Newcastle: LINCOM Europa.

IV Pacific and Australia

William B. McGregor
12 Number in Gooniyandi

Abstract: This paper discusses the category of number in the non-Pama-Nyungan Australian Aboriginal language Gooniyandi. Number marking has various loci in the language, predominantly phrasal, where it is optionally realised by free numerals and/or number marking enclitics. Words are by and large number unspecified; there are just a few exceptions, primarily pronouns, which show number as an inflectional category, and a smallish set of nominal and verbal stems – rarely roots – that are specified for a number value. Nominals themselves are predominantly transnumeral, and unspecified for the count vs. mass distinction. Nominal number is indicated by NP-level enclitics, and also inflectionally in finite verbs, both in the forms of classifiers that occur in all finite verbs, and by number enclitics. Number agreement is at best marginal; instead, the choices of marking or not marking number, and the values assigned, are effectively independent at the various loci. Obligatory number marking in the finite verb apparently plays a significant role in reference management; optional number marking in both the NP and verb seems to be motivated by other considerations, which are not yet adequately understood.

1 Overview

Gooniyandi is an Australian Aboriginal language spoken by a small number of people residing in the southern part of the Kimberley region in north-western Australia. It is a non-Pama-Nyungan language which, along with closely related Bunuba, comprises the Bunuban family (O'Grady, Voegelin *et al.* 1966: 28; Wurm 1972: 123; McGregor 1990a: 1). Like most Australian Aboriginal languages these two languages show a broadly agglutinating profile, though like the majority of non-Pama-Nyungan languages the verb shows significant fusional tendencies.

According to the typological classification of Capell (1940: 244), the Bunuban languages are "prefixing languages without noun classification". That is to say, they have prefixes as well as suffixes, like the majority of non-Pama-Nyungan languages (Pama-Nyungan languages by contrast typically have suffixes only) and no system of noun classes. Nor do they show any other type of nominal classification system such as noun classifiers (in the sense of Allan 1977; Dixon 1982: 217–218), numeral classifiers, or possessive classifiers. However, like many languages of northern Australia there is a system of verb classification in which verbs are classified according to valency, Aktionsart and vectorial configuration (see McGregor 2002). This system is highly grammaticalised in Gooniyandi, more so than in nearby verb classifying languages.

Like the typical Australian language, Gooniyandi has a restricted set of numerals (e.g. Fraser 1890: 236; Nekes and Worms 2006: 110; Dixon 1980: 107–108; Bowern and Zentz 2012):[1] *yoowarni* 'one', *garndiwirri* 'two', *ngarloordoo* 'three, a few', and *garndiwa ~ garndiwangoorroo ~ garndiwangarri* 'many'. Nonetheless, number is a significant feature in the language, and marking of number is exhibited in at least one place in virtually every verbal clause.

Lexical nominals and verbs are mostly unspecified for number: a nominal or verbal lexeme can normally, depending on context, be interpreted as singular or plural in number. There are just a few exceptions where lexical roots or stems are inherently marked as plural or, occasionally, singular. However, personal pronouns, both free and bound, consistently mark number, and distinguish singular from plural. Other grammatical nominal roots and stems (including endophoric, exophoric, and indefinite/interrogative determiners) are, like lexical ones, transnumeral – i.e. neutral with respect to number inflection.

Aside from number marking on pronominal elements and the occasional inherent number markedness of a small number of nominal lexemes, nominal number marking has by and large a phrasal locus, and is non-obligatory. This marking is by free numerals, number marking enclitics to NPs, and/or verbal number enclitics. These marking choices are independent of one another – thus for instance free numerals may occur or not occur with number marking enclitics.

The distribution of number marking is in accordance with the animacy hierarchy (Silverstein 1976): personal pronouns are obligatorily number marked, as are certain human nominal stems (especially terms for social groups and kin categories), while for the rest of the nominal lexicon number marking is optional and decreases in frequency in accordance with reduction in animacy status. Inanimates are rarely number marked. The animacy hierarchy is also relevant to number marking by bound pronominals: the less animate a referent the more likely that it will not show plural marking in the bound pronominal. There are also correlations between number marking and grammatical role. Thus number marking is most strongly associated with NPs serving in participant (i.e. argument) roles, and is less frequent on NPs in other grammatical roles. There are also perhaps different tendencies according to the participant role, whether Actor (roughly "subject"), Undergoer (roughly "object"), or some other role; this is uncertain, however.

---

1 It has sometimes been argued that the typical Australian language possesses no numerals – see in particular Hale (1975), who suggests that the items in Warlpiri and various other Australian languages that translate into English as numerals are in fact indefinite determiners. I have argued at length elsewhere (McGregor 2004, 2007) against Hale's claims, and that Australian languages typically do show a class of numerals. As will emerge as the exposition unfolds, these lexemes share (in Gooniyandi in particular) morphosyntactic and semantic features that motivate grouping them together in a single category. The label 'numeral' is motivated by their common (albeit not exclusive) usage in specification of numerosity, exact or approximate.

Genuine number agreement does not occur either at clause or phrase level; nor in fact does person or any other type of agreement exist. Choices, that is, depend not on linguistic features as such, but rather on features of the referent that the speaker opts to highlight.

2 Pronominal, nominal and verbal number

2.1 Generalities

The vast majority of lexical and grammatical words and morphemes in Gooniyandi are number-neutral, and show no inflectional category of number. The main exceptions are: personal pronouns, which show distinct suppletive inflectional forms for corresponding singular and plural person categories (§ 2.2); derived forms of kin terms and certain reduplicated nominal stems that are inherently plural (see § 2.3); and a few verb roots and derived stems that are restricted to either plurality of events or plural subjects or objects (see § 2.4). Also exceptional are proper names of persons and places, which are apparently inherently singular in reference (see § 2.3).

Number marking is largely confined to phrasal units. Marking of number in NPs is by use of bound grammatical morphemes and lexical numerals, all of which are optional. The bound morphemes, which are phrasal enclitics rather than inflectional morphemes, are discussed in § 3.1; the lexical numerals are discussed in § 3.2. In the VP (by which I mean the inflected verb together with modifying dependents, but excluding arguments) number marking of the verb is by frequency-specifying adverbials and marked forms of numerals (§ 2.4). In addition, bound pronouns in the finite verb cross-reference participant (i.e. argument) roles, and indicate the person and number of the fillers of these roles. (As explained in § 3.3, this is cross-reference not agreement.) This number marking is obligatory, and the person and number categories distinguished are – with one notable exception (see § 2.2) – identical with those distinguished by the free personal pronouns. Clitics in the VP also mark number of NPs in certain roles in the clause (§ 3.4).

2.2 Pronominal number

Unlike nominals, free pronominals in Gooniyandi show separate suppletive roots for singular and plural numbers in each person category. Bound pronominals – which are restricted to finite verbs – show by and large the same person and number categories as the free pronominals, though with some nuances as discussed below.

The person-number system for free pronouns is as shown in Table 1 (see further McGregor 1990a: 170, 1996b). These are the cardinal forms of the pronouns, the

Tab. 1: Gooniyandi free cardinal pronouns.

	Singular	Plural	
1	*nganyi* ({1})	Exclusive	*ngidi ~ ngiyidi* ({1, 2}, {1, 3}, {1, 3, 3}, …)
		Inclusive	*yaadi* ({1, 2, 3}, {1, 2, 3, 3}, …)
2	*nginyji* ({2})	*gidi* ({2, 3}, {2, 3, 3}, …)	
3	*niyi* ({3})	*bidi* ({3, 3}, ({3, 3, 3}, …)	

forms that are generally used when the pronoun serves in a participant or argument role (there are qualifications that we need not go into here). There is also a set of suppletive oblique forms corresponding to each of the cardinal forms, and a paradigm of emphatic forms constructed on the oblique forms; these also show the same person and number category distinctions. These may be appropriately described as personal pronouns in that they are almost exclusively employed in reference to persons and personified beings. The free pronouns are not attested in reference to inanimates, and only rarely in reference to higher order animates other than personified ones.

As can be seen from the extension sets for the pronouns in Table 1, the clusivity distinction made in the first person plural is not the standard inclusive-exclusive one. Whereas in the standard inclusive-exclusive system what is included or excluded is the hearer as a unique individual, in this system what is included/excluded is the category of hearers – the unique addressee plus one or more others. Notice that in this system the first person inclusive category includes a minimum of three referent individuals (the speaker, the hearer and one other), while all other plural categories include a minimum of two. The first person exclusive embraces the speaker-hearer dyad, speaker-other dyad and larger groups of speaker and others (excluding the hearer). See McGregor (1996b) for further discussion of this virtually unique system, found in the neighbouring language Bunuba, and (to the best of my knowledge) just one other language, the Niger-Congo language Yaouré (Hopkins 1986).

The plural cardinal pronouns admit further (optional) number specification by the addition of a number enclitic, either *-yoorroo* DU or *-yarndi* PL (see further §3.1 below). These specify the number of the NP to which they are cliticised, and thus when attached to oblique forms usually apply to the number of the referent of the NP in which the oblique pronominal serves as a dependent, rather than of the referent set of the pronominal itself. Thus *bidi-yoorroo* (3PL.CRD-DU) will specify 'them two', while *birrangi-yoorroo* (3PL.OBL-DU) will normally indicate 'their two Ns'. For all plural cardinal pronominals except *yaadi* 1PL.INC.CRD both *-yoorroo* DU and *-yarndi* PL are regularly attested in speech; as might be expected, *yaadi-yoorroo* is not attested in my Gooniyandi corpus. Some speakers accepted this form when I constructed it, suggesting that it might mean 'we three, speaker, hearer and one other'. (A similar response was encountered by Alan Rumsey for the corresponding form in Bunuba – Alan Rumsey pers.comm.) This could be taken as suggesting that

-yoorroo is a marker of unit augmented ('add one') rather than of duality; however, speakers did not spontaneously use this form, and instead used just *yaadi* 1PL.INC.CRD or *yaadi-yarndi* 1PL.INC.CRD-PL in reference to the triad speaker, hearer and one other.

As remarked above, the bound pronouns in Gooniyandi make by and large the same person and number distinctions as the free pronominals. These bound pronouns are restricted to the finite verb, which can be described by the item-arrangement template shown in (1), where Root indicates the root form of the lexical verb or less frequently of a nominal or adverbial.

(1) Root-(Aspect)-(Mood)-Classifier Complex-(Mood)-(Oblique pronoun)-(Number)

The item labelled Classifier Complex is the only other obligatory item in the finite verb, and is an inflected form of an element I call the classifier, which categorises the lexical verb comprising the Root plus derivational Aspect marker if present; see McGregor (2002: 41–96) for detailed discussion. There is compelling evidence that the Classifier Complex derives historically from inflected verb forms (e.g. McGregor 2002: 93–94). Basically, a classifier morpheme (deriving historically from a verb root) is inflected according to tense as well as to the person and number of the Actor and Undergoer (if there is one). This inflected form is highly fusional in nature, and perhaps not best described in item-arrangement terms: quite abstract underlying forms and rules required, and even then there are exceptional forms. Nonetheless, it is useful in the present context to provide an item-arrangement description of the dominant structure of the classifier complex; this is as shown in (2), where | indicates either a morpheme boundary across which quite significant morphophonemic processes apply or that the items do not always occur in the specified arrangement.

(2) (TNS)|(ACC:PRO)|NOM:PRO|CL

As indicated, there can be up to two pronominal prefixes in a finite verb, one cross-referencing the Actor, the other the Undergoer. This combination is typically preceded by a tense-marking prefix,[2] and followed by a Classifier morpheme that categorises the lexical root according to valency, Aktionsart and vectorial configuration, an abstract configuration of action vectors characterising the event (e.g. downwards motion).

The oblique pronouns in the verb are transparently related to the corresponding free oblique pronouns, and make precisely the same person and number distinc-

---

2 There are many complexities in tense marking that we need not go into here, and sometimes tense is not marked by a segmental morpheme but rather by a fusional form of the Classifier Complex that is distinct from the form found in other tenses. Thus there is no consistent past tense prefix, and for reasons of simplicity I have not indicated past tense in the gloss lines of example sentences.

tions. For the pronominal prefixes in the Classifier Complex things are somewhat muddied by neutralisations and fusions (see McGregor 1990a: 203–212). For those classifiers that take a single pronominal prefix, different forms can be distinguished for each of the person-number combinations, and bound pronoun forms can be identified with relatively few complications. For bivalent classifiers, again different forms exist in most instances for each of the possible person-number configurations for the two actants, though it is more difficult to set up underlying forms for the bound pronouns. And in some combinations, namely when the Undergoer is first person, person of the Agent is neutralised, and only number is specified (McGregor 1990a: 207, 210).

There is one other notable irregularity in the verbal bound pronominals, the existence of a unique speaker-hearer dyadic form *-ngangi* – nowhere else in the language is this category specifically indexed – that optionally cross-references either an Undergoer or an Affected (roughly beneficiary) participant. Thus in some circumstances this form alternates with *-ngirrangi* (1EXC.OBL) 'to us': in (3) both enclitics are possible means of expressing the meanings indicated (though *-ngirrangi* (1EXC.OBL) 'to us' admits other interpretations as well).

(3)   *bala-ji-la*   *wayandi*
     send-IT-1SG.NOM|3SG.ACC|CL fire
     gaj-ba-ngangi
     cut-FUT|2SG.NOM|3SG.ACC|CL-1&2ACC/OBL
     'I sent him$_i$ (saying) "Cut firewood for us (me and you$_i$)!"', 'I sent him$_i$ to cut firewood for us (me and him$_i$).', Or 'I made him$_i$ cut firewood for us (me and him$_i$).'[3]

2.3 Nominal number

Nominals in Gooniyandi can be characterised by the fact that they must either serve in some grammatical role in an NP or as the event-specifying item in a VP (McGregor 1990a: 140–141; 2013a: 232). This part of speech includes grammatical items such as free pronouns, determiners and demonstratives, as well as a diverse range of lexical nominals including terms for human categories, personal names, animals, plants, topographical and environmental features, toponyms, numerals, abstract entities, and qualities. There is no category distinction between nouns and adjectives, and in principle any nominal can serve either as the referential head of an NP or as an attributing element in it. Only free pronouns inflect for number or case; otherwise nominals – including both lexical and grammatical nominals such as demonstra-

---

3 To be completely clear, it should be noted that second and third free translations provide indirect-speech representations of the complement; *-ngangi* in the Gooniyandi original can only refer to the speaker-hearer dyad according to the shifted person deictic system of the represented thought.

tives and determiners – are by and large inflectionally unspecified for number (and case).[4] Thus the majority of nominals as the sole member of an NP admit interpretations as either singular or plural in number: for instance, the nominal *gardiya* 'white person' can be used in reference to a single person or any number of persons; it can also be used generically, as in (13) below. Regular number marking within the NP is, as already mentioned, by means of numerals and/or number marking enclitics (see § 3.1 and § 3.2 below for detailed discussion). The latter are restricted to entities high on the animacy scale, typically persons; numerals are in principle usable for entities of any order of animacy, though their likelihood of use decreases with decrease in animacy.

There are no numeral classifiers in Gooniyandi, and nor is an emic distinction made between mass and count nominals. To be sure, some nominals are normally accorded mass interpretations, others normally count interpretations. However, most nominals that normally show mass interpretations – e.g. *gamba* 'water', *walyarra* 'sand, sugar' and *banda* 'earth, ground' – also admit count interpretations when used in combination with numerals: so many units of the specified item (e.g. units for drinking water or beer). No specification of the unit of measure is ever made.

There are a few exceptional nominal roots and stems that are inherently singular or plural in sense. Thus, personal proper names and toponyms appear to be inherently singular in sense, even though there are occasional instances of namesakes (*narroogoo*). Proper names of persons can take number marking enclitics, as shown by example (20) below. However, these do not (in any attested examples) indicate non-singularity of a set of individuals showing that name (as in *the two Ronnies*), but only the named person plus another or others. In this regard proper names differ from most common human nouns where both interpretations are generally possible.

Just a few nominal roots are inherently plural; the most common is *mawilyi ~ mawoolyi* 'children' – to refer to a single child another lexeme, such as *joogoo* 'child' or *jiginya* 'little, small' must be used; these are number neutral nominals. Another such nominal, also denoting a grouping of persons, is *wilyinggi* 'youths'.

Rather more nominal stems are inherently plural in number. Nominal reduplication – which may be either partial or total – is one derivational process that frequently results in nominal stems that are inherently plural in number. (Nominal reduplication less commonly indicates intensification.) This is perhaps most common for human nominals, where the derived stem specifies a socially relevant group of persons satisfying the specified property, as in e.g. *gambagambayi* 'young boys' (*gambayi* 'young boy'),[5] *boolgawoolga* 'old men' (*boolga* 'old man'), and *barndawarndanyi ~ barndanyibarndanyi* 'old women' (*barndanyi* 'old woman'). A few in-

---

4 Just a single determiner, *yaabja* 'some', appears to be inherently number specified as plural. All other determiners admit reference to either singular or larger groups.

5 Note that the unreduplicated forms are number-neutral despite the glosses.

stances of reduplication of inanimate nominals are attested, including *jarranjarandi* 'twigs' (*jarrandi* 'twig') and *dagoorladagoorla* 'covered in depressions' (*dagoorla* 'depression, hole').

These inherently plural roots and stems need not necessarily be used in reference to unified groups, and like other nominals admit generic interpretations. Moreover, they may be used both referentially and attributively.

Kin terms admit additional derivational processes, sometimes in combination with reduplication, that result in stems that are inherently plural in number. As in many Australian languages (see e.g. Merlan and Heath 1982; Evans 2006), there is a dyadic kinterm suffix – *-langi* in Gooniyandi – that derives from a kin term another that specifies a pair of individuals in the specified kin relation to one another. Thus for instance *garingi-langi* (wife-DY) specifies a man-wife pair, *goornda-langi* (cousin-DY), a pair who are related as cousins. Not all kin terms are attested with *-langi* DY; there are also a few other nominals that admit this suffix, e.g. *jaliji-langi* (friend-DY), a pair of friends. As might be expected given that the dyadic suffix is a derivational one, there are many irregularities and complexities in the system; we need not go further into them here – see McGregor (1996a). Of course, kin terms also admit number enclitics. With the dual enclitic the duality is of the referent of the NP, so that e.g. *garingi-yoorroo* may mean 'two wives' or 'two individuals, including someone's wife': the referents may or may not be in the specified kin relation to one another.

Gooniyandi shows a further elaboration in kin term derivation: polyadic kin terms, which specify groups of kin that are either related to one another in a particular kin relation, or are in that relation to some other individual as propositus. These are formed in various ways: by the reduplication of a dyadic kin term; by the partial reduplication of a kin term together with the dyadic suffix; and by the suffixing of the reduplicated form *-langi-ga-langi* (DY-*ga*-DY) to a kin term. For instance, *marni-marni-langi* (older:sister-older:sister-DY) indicates a group of individuals related as siblings to one another, and *marni-langi-ga-langi* (older:sister-DY-*ga*-DY) indicates a set of individuals who are all related as siblings to some propositus (note the potential overlap in reference of these terms). Again see McGregor (1996a) for discussion of irregularities and complexities.

The Gooniyandi social universe is divided into eight subsections, to one of which groups everyone belongs (see McGregor 1990a: 14–16). Like other nominals, the subsection terms are not number specified. However, they admit the suffix *-warnoo* which indicates the entirety of members of the subsection in the relevant context. Thus, *jagarra-warnoo* specifies all of the contextually relevant members of the *jagarra* subsection.

Similarly, place names admit the suffix *-warrawarra* indicating everyone associated with that particular place, for instance, everyone who lives there, everyone with that place as their conception cite, or whatever. A corresponding suffix *-waangoo* specifies a single person from the specified place.

2.4 Verbal number

Verbs are like nominals in that they are number neutral. Depending on context, a lexical verb admits interpretations as specifying a single event or a number of events, the latter including habitual and generic interpretations. For instance, (4) admits the single event interpretation (specifying a single instance of drinking at some point in the past) and the habitual interpretation (which in the past is normally understood pragmatically to refer to the habitual consumption of a particular liquid such as beer). Generic interpretations are also possible, where the reference is not to any individual instance or instances of the event type, but to non-specific occurrences, as in (5), from a text describing traditional marriage practices. These remarks apply to other tenses as well (though generic interpretations are less likely in future tensed verbs). (See further McGregor 1990a: 516–523.)

(4) *gamba ngoorloog-nga*
    water   drink-1SG.NOM|3SG.ACC|CL
    'I drank water', 'I used to drink beer.'

(5) *goornboo liya nyoon-birra   goornboo-yarndi-ga*
    woman   west rub-3PL.NOM|3SG.ACC|CL woman-PL-ERG
    'To the west the women would rub the woman (with ochre).'

Explicit specification of number is possible for verbs by use of frequency adverbials and particles. There are three frequency adverbials that are perspicuously constructed from numerals, *yoowarningarri* 'once', *garndiwirrja* 'twice', and *ngarloorrja* 'thrice'. There is no corresponding frequency adverbial for the numeral *garndiwa ~ garndiwangoorroo* 'many'; instead, the particle *ngarrarni* 'always, all the time, many times' is used. Examples (6) and (7) illustrate the usage of these items.

(6) *yoowarningarri gard-boowoo*
    once   hit-FUT|2SG.NOM|3SG.ACC|CL
    'Hit him once.'

(7) *ngarrarni girragirra-yi/   yiliyilib-ji   ngarrarni/*
    always   run-3SG.NOM|CL sneak-3SG.NOM|CL always
    'He was always running away, he was always escaping.'

The frequency adverbials are also used somewhat more generally in the specification of temporal duration, as shown by examples (8) and (9). In these examples days are counted by their use. Whereas in (8) the event of lying down might be counted, this cannot apply to (9), where just one instance of the speaker's absence

is being referred to, and its duration specified. Occasionally the frequency adverbials are further specified by the nominal *riwi* 'place, country, day'.

(8) *niyaji-ya bagi-yirri   garndiwirrja/*
   this-LOC lie-1.EXC.PL.NOM|CL twice
   'We camped there for two days.'

(9) *garrig-goowa-wingirni   ngarloorrja*
   absent-PRG-FUT|1SG.NOM|CL|POT thrice
   'I could be away for three days.'

Verb roots that specify telic events admit a derived stem that indicates an iterated sequence of events of that type, the new stem being classified as atelic. For instance, from *gard-* 'hit' we have *gard-bi-* 'hit repeatedly, belt, beat', and from *ngang-* 'give' there is the derived *ngang-ji-* 'give repeatedly, e.g. feed'. The derived form thus specifies both plurality of the events, and that the repeated instances belong together to form a new event type. Like the dual and plural markers in the NP, the iterative thus specifies collectivity as well as number.

Another derivational process with a similar semantic effect is reduplication of a verb root. For instance *gaj-* 'cut' reduplicates as *gajgaj-* 'cut up, chop up', *mird-* 'tie up' reduplicates as *mirnmird-* 'tie up repeatedly', and so forth. But reduplication of verbs sometimes has other semantic effects. For some verbs, what is represented is repetition of action distributed over a range of different participants, as in e.g. *lalbalalbag-* 'split all over' (*lalbag-* 'split'), *waroongwaroong-* 'fly overhead (of a number of birds)' (*waroong-* 'fly') *lalaj-* 'lie all about (as of a number of footprints of different people)' (*laj-* 'lie on the ground (of footprints)'). Reduplication can also indicate action over an extended period of time, as in e.g. the reduplication of *ngarag-* 'work' to *ngarangarag-* 'work on something for a long period of time'. However, it seems likely that such mass interpretations are only etically distinct, and that the same reduplications can also be used in reference to repeated actions by the same or multiple entities.

Just a few verbs are number restricted in the sense that they require a plural argument. (Of course, it must be acknowledged that evidence is restricted.) The majority of these are intransitive verbs that require a plural Actor, as illustrated by the following: *balbirr-* 'return (of many people)', *barlwarl-* 'lie about', *baward-* 'climb up', *birrabirra-* 'scatter, move in different directions', *gawoorrij-* 'pass away or die, of many people', and *girnaj-* 'sit around as a group'. There are just a couple of known transitive verbs of this type, and the plural argument may be either the Agent, as in *garrawan-* 'gather around, surround', or the Undergoer, as in *maroorr-* 'muster, draw in to a group, gather, heap up things, foster, bring up someone else's children'.

There are a couple of other lexico-grammatical phenomena in Gooniyandi that relate more or less closely to the expression of number and/or quantity in the

verbal domain. For instance, there is a particle *ngambirri* 'again' and an enclitic *-nyali* REP both of which can be used to indicate another instance of an event of the type already mentioned. An example is the verb form *dij-nyal-aa* (snap-REP-PRS|3SG.NOM|3SG.ACC|CL) meaning 'he is snapping it again', indicating the occurrence of another instance of the specified type of event.

3 Agreement and the syntax of number

Agreement plays at best a very minor role in Gooniyandi grammar. Within the NP there is no agreement as such in terms of the traditional categories of number, case, and gender. As mentioned already number marking is optional in non-pronominal NPs, and is expressed by means of NP-level clitics and/or numerals. A given non-pronominal NP may express number by either or both of these means, though very often number goes unexpressed. If both representations occur in an NP, they are of course consistent: the numeral *garndiwirri* 'two' will only occur with the dual enclitic *-yoorroo* within the same NP. In NPs with pronominal heads dual and plural marking of that head is possible only for the plural members of each person category; as we have seen, the dual marker is not attested with the inclusive pronoun *yaadi*, though many speakers do not reject it. Nor are the number enclitics added to singular pronominal forms in coordinate constructions such as 'me and him' or 'you and her' – instead, a non-singular pronoun is employed with the relevant number enclitic, as in 'we-two including him' and 'you-two including her' (see further §3.1 and discussion of examples (20) and (21) below).

Number marking in Gooniyandi plays out in a range of syntactic positions. These do not show agreement. Rather in each position separate systems of numerosity are invoked, which differ in slight but significant ways from one another. It is not the number specification of a syntagmatically related item that determines the number marking in a particular locus of number marking. Instead, what is relevant is the referent and the number marking paradigm at that particular locus. For instance, numeral modification, coordination, and so on in the NP do not predict number marking in the verb. Animacy is significant to number marking, such that it is most likely for humans, least likely for inanimates. For humans, number marking in the cross-referencing pronominals in the verb is highly consistent; for inanimates it is infrequent. For other degrees of animacy it is somewhere in between. If plural marking is used of an inanimate entity, it is most likely to be found in the NP, specified by a numeral. The locus of numerals and number marking in NPs (but not VPs) is somewhat variable, but different constructions – and hence meanings – are associated with the different loci.

3.1 NP number marking enclitics

Gooniyandi has two bound number markers *-yoorroo ~ -yirri* DU and *-yarndi* PL; the initial palatal glide hardens to the palatal stop when preceded by a stop or nasal. The two dual allomorphs *-yoorroo* and *-yirri* appear to be in free variation; they are not phonologically conditioned. Different speakers prefer one or the other; perhaps they were once dialectal variants, but at the time I did my fieldwork no dialectal correspondence was evident.

These two number markers are phrase-level enclitics that indicate the number of the NP they are attached to, as shown by examples (10) and (11). Like case marking enclitics they enter into syntagmatic relations with NPs rather than with words. Both types are bound phrase level morphemes that are cliticised to the focal word of an NP, the word that carries the most significant or unpredictable information: in examples (10) and (11), this is the main lexical item in the NP. Notice in (10) that in the first line it is the type specifications of the referent that are focalised by the presence of the dual marker on the two lexical expressions; in the second line the marker occurs instead on the first nominal element of the NP, which specifies a new feature of the referent, the wearing of pearl shell pendants.

(10) *yoowayi/ niyaji/ yangbala-yirri   yoowooloo-yirri/*
     yes   this   young:person-DU man-DU
     jagooli-ngarri-yirri   yoowooloo garndiwirri/
     pearl:shell:pendant-COM-DU man   two
     jagooli-ngarri/
     pearl:shell:pendant-COM
     'Yes, (there were) these two young men, two men with pearl shell pendants, with pearl shell pendants.'

(11) *girnaj-birri   garndoowoongoorroo yoowooloo-yarndi maa*
     sit:around-3PL.NOM|CL many   man-PL   meat
     garnanganyja ngab-birra
     emu   eat-3PL.NOM|3SG.ACC|CL
     'A lot of men sat around together eating emu meat.'
     (from the draft dictionary of Gooniyandi, Anon 2004)

Number and case-marking enclitics, as expected from the above remarks, typically cooccur on the same word. As might be predicted given that the number markers indicate properties of the NP referent while case markers specify relations NPs enter into, number enclitics usually precede case marking enclitics. The only exceptions to this generalisation occur in circumstances in which the case marker indexes NP-internal (rather than clausal) grammatical relations, as in (10). That is, when the number marker pertains to the number of the outermost NP in an embedded

construction. In sum, Gooniyandi differs from the putative Australian norm, where number markers are allegedly stem forming suffixes (Dixon 1980: 323), and case markers inflectional suffixes (Dixon 1980: 292).

Number marking enclitics usually occur on NPs that refer to human beings or personified animals (e.g. in mythology), and only rarely on NPs that refer to higher order animates; they almost never occur on NPs referring to lower order animates or inanimates. Moreover, the NPs that are marked by number marking enclitics almost always serve in participant (i.e. argument) roles, as can be seen from an inspection of the examples given in this section.

As is the case for case marking enclitics, more than one instance of a number marking enclitic is occasionally found on what might appear to be an NP, as in examples (10) (first line) and (12). However, as I have argued elsewhere, the effect of attaching more than one enclitic to a syntagm that might constitute an NP is to split it up into two separate NPs (McGregor 1989b; see also McGregor 1997). Thus in (12) the enclitics could also occur on just *yoowooloo* 'man' or just *goornboo* 'woman', in which cases the most appropriate translation would be 'He looked for the men and women' (with different foci on *men* or *women*).

(12) *moow-nga-wirrangi*   *yoowooloo-yarndi-yoo*
     look-3SG.NOM|3SG.ACC|CL-3PL.OBL man-PL-DAT
     goornboo-yarndi-yoo
     woman-PL-DAT
     'He looked for the men and for the women.'

Both dual (specifying two) and plural (specifying more than two) enclitics are collective in sense: they indicate as well as number also the sense that the referents belong together as a unified group. Thus in (10) the two referents are treated as a single set, a pair of initiates, while in (11) the men are sitting around and eating together as a group. The number marking enclitics are not used in expression of universals or generalities, where the referents need not belong to a coherent set, as in e.g. *men smoke Drum*. Thus the plural enclitic is not used in (13), where the NP *gardiya* 'white person' is interpreted generically but not collectively. There are some apparent exceptions, e.g. (14). However, the free English translation is misleading in that it does not make it clear that the type of hunting being referred to is group hunting by ambush. While the NP is being used generically, the number marker indicates that it is a coherent group of these unspecified persons that are acting together.

(13) *ngoorndoongoorni-ya/ gardiya-ngga*   *goowaj-goorra*
     whatchamacallit-LOC white:person-ERG name-PRS|3PL.NOM|3SG.ACC|CL
     binibig yaad/
     Spinifex Yard
     'At what's it called; white people call it Spinifex Yard.'

(14) *wanyjirri joog-gila-wirra*   *ngamoo yoowooloo-yarndi-ga*
      kangaroo hunt-FCT-3PL.NOM|3SG.ACC|CL before   man-PL-ERG
      jinali-ngarri-ngga
      spear-COM-ERG
      'A long time ago, men used to hunt kangaroos with spears.'

Given that the number marking enclitics specify a collective sense, it is not surprising that they are not infrequently found in the context of more complex syntagms than simple NPs, where membership of the collective is more precisely specified. They are sometimes found on NPs that encapsulate lists of conjoined NPs, effectively summing them up, and indicating that the referent group is a unified one. This is shown by (15) and (16). In (15) the first two NPs, personal names, are conjoined to form a single group referred to in the following two NPs which are marked as dual. In (16) the first NP is the plural pronoun *gidi* 2PL.CRD; the composition of the referent group is spelt out in the following list of NPs that are conjoined together. Note that in the conjunction of the NPs no marker of the conjunction is employed.

(15) *goodoorrja/ nyimangoo/ maroowa-yoorroo/ niyaji-yoorroo-ngga*
      Goodoorrja Nyimangoo murderer-DU   this-DU-ERG
      gard-boondoorni/
      hit-3PL.NOM|3PL.ACC|CL
      'Goodoorrja and Nyimangoo, two murderers, they two killed them.'

(16) *gidi-yarndi lambadi*   *ngaanggi garingi ngaanggi nginyji*
      2PL.CRD-PL wife's:father 2SG.OBL wife   2SG.OBL 2SG.CRD
      'You lot: your father-in-law, your wife, and you.'

Such lists need not be complete: for one reason or another, only partial specification of the membership of the collective might be given. Often this is because the other member or members are predictable. In some instances the least predictable of the referents only is specified, as in (17), where of course the speaker must be one of the referents of the first person group. And sometimes just the most significant of the collective are specified, as in (18), where reference is made to three crocodiles that have been killed and cooked in trenches. Specific reference is made in the first NP to the two big ones, providing new information about the crocodiles – that two were big. No mention is made of the less significant one, the smaller crocodile.

(17) *niyi-nhingi/ niyi-nhingi barn-jirri-yi*   *ngidi/*
      that-ABL   that-ABL   return-1PL.EXC.NOM|CL-DL 1PL.EXC.CRD
      limba-yoorroo/
      policeman-DU
      'Then we went back, me and the policeman.'

(18) *nyamani-yoorroo/ ngarloodoo niyi-yarndi gird-birrini/*
    big-DU   three   that-PL   leave-3PL.NOM|3SG.ACC|CL
    goorrgoo-ya/
    hole-LOC
    'The three of them, including the two big ones, they left in the hole.'

As shown by (19), the dual enclitic sometimes appears to function as a conjunction. This interpretation however derives from the fact that it marks the number of the NP as dual, and that the two referents form a collective. Put in other words, while the nominals in this example are conjoined, it is not their conjunction but number that is marked by the enclitic.

(19) *mird-jirrrimi-yi/*   *goornboo-yoorroo:: yoowooloo/*
    tie-1PL.EXC.NOM|3PL.ACC|CL-DL woman-DU   man
    'We tied the two of them up, the woman and the man.'

Furthermore, as is the case generally in the language, ellipsis of nominals and NPs is not uncommon, and the other member of the conjoined set need not be specified, as in (17) (where the speaker is not listed in a separate NP) and (20). The occurrence of the dual enclitic on *ngoorndoongoornoo* 'whatchamacallit' in (20) indicates that the Undergoer referent set is two in cardinality and the bound pronominals in the verb indicate that it is a first person exclusive group; the second member can only be the speaker.

(20) *lamaj-jindi-yi*   *garndiwirri/*
    pick:up-1PL.EXC.ACC|3SG.NOM|CL-DL two
    ngoorndoongoornoo-yoorroo/ wiyirri-yoorroo/ wiyirroo/ jawalyi/
    whatchamacallit-DU   Wiyirroo-DU   Wiyirroo jawalyi
    lamaj-jindi-yi/
    pick:up-1PL.EXC.ACC|3SG.NOM|CL-DL
    'He picked up two of us, me and whatchamacallit, Wiyirroo, the Jawalyi man.'

Examples such as (20) are often referred to as inclusory constructions. However, for Gooniyandi there is no reason to postulate a separate construction type: the expressions can be interpreted as elliptical, with the second member of the pair left unspecified (though inferable). Even examples like (21), where the inclusory meaning comes to the forefront so to say, do not argue for a separate inclusory construction. Here the most plausible analysis is that there is a discontinuous NP *gardiya-yoorroo goornboo* 'white man and woman', and that the discontinuity is motivated by considerations of information delivery, to make the woman the unmarked focus of the first clause; the previous material provides a characterisation of that woman in terms of a proposition that is presented as a fact (see further McGregor 1988).

(21) *gardiya-yoorroo   bagi-la-woorroo-yoo   goornboo niyaji-ngga*
     white:person-DU lie-FCT-PRS|3PL.NOM|CL-DL woman   this-ERG
     ngang-ngindi   nganyi
     give-1SG.ACC|2SG.NOM|CL 1SG.CRD
     'The woman who lives with the white man gave it to me.'

Only the dual enclitic is attested in this putative conjunctive usage, that is, on one conjoined nominal in an NP. Whether or not the plural enclitic can be used in a comparable way is uncertain. There are no examples in my corpus, though I have not tested such constructions for acceptability. In usage, however, it is clear that when the referent group is larger than two, other strategies are favoured in the (full or partial) enumeration of group membership. In particular, there is a clear preference to use an expression that encapsulates the entire set, as in examples (16) and (22), and to mark this expression (rather than any of the conjuncts) with the plural enclitic.

(22) *niyaji-ngga ward-ji-wirrangi/   doodoo-ji-wirrarni/   maadi/*
     this-ERG   go-3SG.NOM|CL-3PL.OBL shiver-IT-3PL.NOM|CL cold
     yoowooloo-yarndi-ga/ boolgawoolga-ngga barndanyibarndanyi-ngga/
     man-PL-ERG   old:men-ERG   old:women-ERG
     'He went up to them who were shivering from the cold, all the people, the old men and the old women.'

Number marking enclitics are optional in all circumstances bar one. NPs specifying sets of two or more individuals that act as a unit may or may not be marked by a dual or plural enclitic. The NP without one of these enclitics is unspecified for both number and collectivity. The exception is that the dual enclitic obligatorily occurs on a single nominal that serves as the only overt member of a conjunction, as in examples such as (17) and (20): omission of the enclitic *-yoorroo* DU is impossible without a complete change in meaning and grammatical structure.

3.2 Number marking by numerals

Numerals occur in two main places in the Gooniyandi NP, in the role of Quantifier or Qualifier in the formula of (23),[6] which describes NPs as sequences of grammatical roles served by the component words (see further McGregor 1990a: 253, 1992). It will be observed that all NPs have an Entity role, a role that is served by an item that provides the major specification of the referent; all other roles are optional,

---

6 Numerals can also serve in the Deictic role, where they effectively act as indefinite determiners (McGregor 1990a: 258–259).

and may or may not be present in a given NP. Ellipsis of the filler of any role is also possible, including of Entity. Thus NPs consisting of units serving in effectively any combination of these roles is grammatically acceptable.

(23) (Deictic) (Quantifier) (Classifier) Entity (Qualifier)

For lexical NPs, that is, NPs with lexical nominals in the Entity role, numerals most often serve in the Quantifier role, and so precede the lexical nominal (if present). By contrast, for pronominal NPs numerals normally serve as Qualifiers, and follow the pronoun. I have argued that this difference in behaviour is a consequence of the fact that items in pre-Entity roles serve to modify the reference of the Entity nominal, while post-Entity items modify the referent. For nominals, numerals typically select sets of entities of the given lexical class, whereas for pronominals the referent set is normally already established, and the number is attributed of the referent set. The reverse orders are possible, but give rise to marked meanings, expressed in English by such expressions as 'the men who are two in number' and 'the two of them'. To illustrate the semantic difference associated with the different orderings of numerals, consider the minimal pairs (24) and (25). In (24) the numeral serves to select a particular number of sons as referents (thus modifying the reference of the phrase), while in (25) the number is attributed of the sons, modifying the referent, specifying a characteristic of an already known set, namely that they are two in number. (See McGregor 1990a: 270–272 for further discussion of the contrast.)

(24) *ngarragi garndiwirri ngaloowinyi*
      1SG.OBL two   son
      'my two sons'

(25) *ngarragi ngaloowinyi garndiwirri*
      1SG.OBL son   two
      'my sons who are two in number, the two of my sons, both of my sons'

All numerals can also be used attributively in clauses as well as in NPs. For instance, they can be used in verbless clauses to specify the numerosity of a set, as shown by (26), where the numeral attributes numerosity to the group referred to by *niyaji* 'this', established in the prior text as a group of sick people. More often, numerals express secondary predicates of a participant in a verbal clause, as in example (27), where the numeral attributes a quantity to the Agent as they are performing the activity. (In keeping with remarks above, the marking of both *nganyi* 1SG.CRD and *yoowarni* 'one' by the ERG enclitic indicates that the two words represent separate phrasal units.)

(26) *niyaji garndiwangoorroo/*
    this   many
    'They were many.'

(27) *igi marriyali   ngirnda nyamani girli gand nganyi-ngga yoowarni-ngga*
    no wife's:mother this   big   same can't 1SG.CRD-ERG one-ERG
    ngab-bila/
    eat-FUT|1SG.NOM|3SG.ACC|CL
    'No, mother-in-law, it's too big. I can't eat it alone.'

See McGregor (2005) for further discussion of numerals as secondary predicates in Australian languages.

Another nearby grammatical phenomenon is illustrated by (28), where the repeated numeral specifies the sequential sense 'number by number', expressing what I refer to as iterated involvement.

(28) *yoowarni yoowarni/ dardigirr-mi-wirrra/*
    one   one   learn-IT-3PL.NOM|3PL.ACC|CL
    'They learnt English word by word.' Or 'They learnt English one word at a time.' Or 'They learnt English little by little.'

Occasionally a numeral attributing on a clausal participant appears in a parenthetical expression, as illustrated in the following two examples.

(29) *niyi-ngga boorroonggoo ward-bina   yoowooloo niyaji/*
    that-ERG north   go-3SG.NOM|3PL.ACC|CL man   this
    garndiwangoorroo/
    many
    'He took these men north, all of them.'

(30) *bagi-yirrani   yoowarni-ngarri-nyali riwi/ riwi/ mangarri*
    lie-1EXC.NOM|CL one-COM-REP   day day not
    garndiwangoorroo/
    many
    'We camped there for one day, no more.'

3.3 Number marking by bound pronominals in the verb

Turning now to the verb, we have already seen that although the system of person and number marked by bound pronouns is by and large identical with that of free pronouns, strictly speaking the verb does not show agreement with NPs serving in

participant roles. Rather, the bound pronominals in the verb cross-reference the NPs filling these roles – or more accurately their referents. Thus a number-neutral NP serving in a participant role may be indexed in the verb by either a singular or a plural bound pronominal. Moreover, the exceptional bound pronoun *-ngangi* 1&2ACC/OBL does not agree with any person, number or case feature marked by a free pronoun that it cross-references. Bound pronominals also distinguish a different case system from the free NPs (including pronominal ones), a nominative-accusative-oblique system, in contrast with an ergative-unmarked-dative system marked by the case enclitics.

Every finite verb can be analysed as containing between one and three bound pronouns. These index fillers of participant roles according to their person and number, regardless of animacy. In contrast with the free pronouns which are virtually never used in reference to non-human entities, nominative and accusative bound pronouns frequently index non-human, even inanimate entities; oblique bound pronouns do also, albeit rarely (McGregor 1990a: 332).[7] However, the likelihood of use of a plural bound pronoun for a non-singular set of entities decreases with lesser animacy. Plural sets of higher order non-human animates are not infrequently cross-referenced by plural third person bound pronouns, though alternatively they are quite often cross-referenced by singular ones. For instance, in a text describing a group of men killing three bullocks the narrator describes this event in two distinct sentences, separated by another sentence. In the first, (31), the cattle are indexed by a plural bound pronoun, in the second, (32), by a singular one.

(31) *ngarloodoo gard-boondini-rri/ ngarloodoo/ booloomani/*
three hit-3PL.NOM|3PL.ACC|CL-PA three bullock
'They killed three bullocks.'

(32) *nyag-birrini/ jinali-ngarri-ngga/*
pierce-3PL.NOM|3SG.ACC|CL spear-COM-ERG
'They speared (the bullocks) with spears.'

For plural inanimates, the indexing bound pronoun is almost always a singular one, as shown by (33), which context makes clear is referring to the making of many spears, not just a single one. However, this is not a rule, but a strong tendency, as shown by (34), where the plural third person pronoun in the second verb cross-references the rocks.

---

7 One other difference between free and bound pronouns is that free pronouns are never used non-referentially like the English pronouns *it* (in *it is raining*) and *they* (as in e.g. *if you don't pay your taxes they will get you*). Bound pronouns in Gooniyandi are sometimes used non-referentially, as for instance is the case for the third person plural nominative prefix in an impersonal construction (see further McGregor 1990b). A more complicated case of a putative zero non-referential bound pronoun is discussed in McGregor (2003: 97–100).

(33) *jinali mirn-mird-jirra:: wili/*
spear tie-tie-1EXC.NOM|3SG.ACC|CL finish
'We tied up spears.'

(34) *booladi ngaarri wara-woorra-arri*
dry   rock   stand-FUT|3PL.NOM|CL-PA
'Dry rocks will be standing around.'

3.4 Number marking enclitics on the verb

As shown in the item-arrangement formula (1), finite verbs admit number marking enclitics in their final order-class. There are two such enclitics, *-yi ~ -yoo* DL and *-rri ~ -rroo ~ -warri ~ -arri* PA. The first specifies the duality of the filler of some participant role – in other words, it specifies more precisely the number of one of the bound pronouns in the finite verb – while the second specifies that the number is a few, not many. The PA category overlaps with but does not coincide with the category marked by *-yarndi* PL in the NP: the latter as we have seen can specify any number greater than two, whereas the former specifies that the number is several. If the argument role indexed by the PA enclitic is filled by an NP with number specification, this is normally by the numeral *ngarloodoo* 'three, a few', as in (31) and (35). It may however use the numeral *garndiwangoorroo* 'many', as shown by (36). In this instance, it turns out that the group of people who have been gathered together are three in number, and the numeral *garndiwangoorroo* 'many' is being used to emphasise that the entirety of the set was involved, rather than that their numerosity was more than a few.

(35) *ngarloodoo yoowooloo bagi-yirri-rri/*
three   Aborigine lie-1EXC.NOM|CL-PA
'Three of us Aborigines camped (there).'

(36) *garndiwangoorroo-nyali maroorr-loonbini-rri/*
many-REP   muster-1SG.NOM|3PL.ACC|CL-PA
'I gathered them all together.'

One final qualification needs to be added concerning the use of the verbal number enclitics for the first person non-singular categories. As expected, for the inclusive category, the dual enclitic never occurs, only the paucal. For the first person exclusive both dual and paucal markers are attested, but the dual occurs only when the first person group is a dyad involving the speaker and someone other than the hearer. When it is the speaker-hearer dyad, the dual enclitic may not be used, and only the bound pronoun *-ngangi* (see § 2.2 above) occurs.

4 Semantics and discourse

4.1 Nominal meaning and number marking

There seems to be no emic contrast between mass and count nominals in Gooniyandi, though of course certain nominals are most likely to be used in count contexts, others in mass contexts. Nonetheless, lexical nominals are largely unspecified on the mass-count dimension, and many are used both in the mass sense and in the count sense. In principle any nominal can be used with number specification, e.g. occur in an NP with a numeral, to indicate a number of instances of a given category, construed as something countable. Measure words and classifiers are not employed in such circumstances. Quantity specifiers are also unspecified in mass vs. count terms. Numerals like *yoowarni* 'one' and *garndiwangoorroo ~ garndiwa* 'many' admit mass interpretations 'a little' and 'much', respectively, in certain contexts. Conversely, mass quantity specifiers are sometimes employed in count senses, so that *jiginya* 'little' and *nyamani* 'big' admit, given the appropriate context, count interpretations 'few' and 'many' respectively. The rather marked mass quantifier *ngarrja* 'little, not much' is particularly prone to be used in the count sense 'few'.

Nominals in Gooniyandi are also unspecified on the type-token dimension: they can in principle be used in reference to both tokens of a particular type and to the categories themselves, conceived of as units. The counting of types in Gooniyandi uses the morphological devices employed for counting tokens; no special marking is used. To illustrate this consider subsection terms. In narratives these are typically employed in reference to tokens, members of the subsection. Occasionally in expository texts these nominals are used in reference to types. In one exposition about the Gooniyandi subsections number marking was in accordance with the subsections as types. Thus where reference was made to a single subsection as a category, number was marked as singular, as shown in (37), where two subsections are each separately cross-referenced by singular bound pronouns in the verb, though of course each has many members.

(37) *naangari-ngga ward-ga*   *jangala/ joogoo/ ngirrangi/*
      naangari-ERG go-PRS|3SG.NOM|3SG.ACC|CL jangala child   1EXC.OBL
      ngaliwinyi/
      son
      'Naangari has Jangala sons, our children.

By contrast, (38) and (39) invoke reference to two subsections together as groups. In (38) the kin dyad and the pronoun both specify duality, and refer to the two subsections, not their members. In (39) it is the cross-referencing pronouns in the verb and the number enclitic on it that mark that two subsections are referred to.

(38) *marni-langi   bidi-yirri/*
   older:sister-DY 3PL.CRD-DU
   'They (Jawandi and Nyawana) are brother and sister.'

(39) *nangala/ jangala/ ngirrangi jaminyi   ward-goorra-yi/*
   nangala jangala 1EXC.OBL grandfather go-PRS|3PL.NOM|3SG.ACC|CL-DL
   jawalyi/ nd nyawajarri/ jaminyi/   ngirrangi
   jawalyi and nyawajarri 1EXC.OBL grandfather
   'Nangala and Jangala have children who are our grandchildren, Jawalyi and Nyawajarri.'

There are instances in which number marking is so to say intermediate between marking of tokens and marking of types, where it applies to speaker-construed sets comprising more than one member. Consider (40). Here it is clear that four people are being referred to in the set of people taken. But the number marking on *niyaji* 'this' is dual. I do not believe that this is a speech error. Rather, it seems reasonable to presume that *niyaji-yoorroo* 'this two' is being used to refer to the two groups, these being the two men and the two women. (As to the paucal marking on the verb, it is not clear whether this applies to the Agent, the first person group, or the Undergoer, the four others.)

(40) *garndiwirri/ garndiwirri goornboo:: garndiwirri yoowooloo/*
   two   two   woman   two   man
   niyaji-yoorroo-moowa doow-yirrra-rri/
   this-DU-ON   get-1EXC.NOM|3PL.ACC|CL-PA
   'Two women and two men, that's all we got.'

4.2 Usage of number marking in discourse

As is typical of Aboriginal Australia, the Gooniyandi traditionally showed respect and avoidance towards in-laws, manifestations of which are evident in the tenor of interpersonal behaviour, including speech behaviour and other communicative phenomena. A special speech register was employed in Gooniyandi in communication with and about key 'close' in-laws, most notably between a man and his actual wife's mother, to a lesser extent between a man and classificatory wife's mothers, and their spouses, wife's fathers – see McGregor (1989a, 1990a: 16–20). One characteristic of this register is the use of plural pronouns in reference to and address of single individuals in 'close' avoidance relations to a speaker. Thus in address *gidi* 2PL.CRD would be used instead of *nginyji* 2SG.CRD, and in reference, *bidi* 3PL.CRD rather than *niyi* 3SG.CRD; the speaker-hearer dyad would be denoted by *yaadi*

1INC.CRD rather than *ngiyidi* 1EXC.CRD. This usage also applies to oblique forms of the free pronouns and to bound pronouns in the finite verb.[8]

Free pronouns in Gooniyandi are normally used with specific and definite reference; they are not attested in generic usage. As far as I am aware, this meaning is normally expressed by indefinite/interrogative determiners, possibly together with a clitic such as *-wirri* UNKN or *-ngarraya* TOO. Thus instead of a second person pronoun in a clause such as *you never know*, the indefinite determiner *ngoorndoo-ngarraya* 'someone-TOO' (roughly, 'anyone') would be used in Gooniyandi.

We have already seen that ordinary nominals can be used in either singular or plural reference, or generically; thus depending on context *yoowooloo* 'man' can be used to mean 'one man', 'more than one man', or 'men in general'. When used generically, cross-referencing in the verb is normally plural if the referent is human (see examples (13) and (14) above), singular or plural if it is a higher order animate (in (14) *wanyjirri* 'kangaroo' is used generically, and is cross-referenced by a singular bound pronoun), and normally singular if it is inanimate (see examples (4) and (41), where *gamba* 'water, beer' and *yawoorna* 'type of ceremonial gear' are used generically). It will be recalled that number marking enclitics on NPs specify collectives, and are only used in generic reference where the generically specified entities behave together as a group.

(41) *ngamoo daj-gila-widi/*   *yawoorna/*
     before   wear-FCT-3PL.NOM|3SG.ACC|CL ceremonial:gear
     'Before (in the old days), they used to wear *yawoorna* ceremonial gear.'

Person and number marking in bound pronouns in the finite verb has obvious grammatical and discourse usages in indexing the fillers of grammatical roles, and tracking them across finite clauses. Information on grammatical roles served by an NP is only partly conveyed on the NP itself. Moreover, ordinary Gooniyandi speech is highly elliptical: in a small corpus,[9] just over a quarter of the participant roles are realised by NPs; the overwhelming majority have no overt expression except in the bound pronominals in the verb. Thus the bulk of the information about the grammatical roles is conveyed by bound pronouns – and the inferences that can be made from them in context concerning the roles that would have been marked on the NPs if they were there.

Number marking in the bound pronouns is highly consistent for human participants, and presumably facilitates reference tracking and the interpretation of gram-

---

8 Another prototypical avoidance relation in Australian Aboriginal societies is between a man and his sister. Most likely the same skewing in pronoun usage existed in the speech behaviour of adult opposite sex siblings; unfortunately, I have no clear evidence of this.

9 This comprises some fifty-one mainly monologic texts (forty-nine narratives, two expositions) related by half a dozen speakers. In total this comprises 6333 sentences, and 3969 verbal clauses.

matical roles. For lower order animates and inanimates, less is explicit and inferencing carries a heavier burden. Elsewhere in the finite verb participant number is optionally marked by dual and paucal enclitics. Their usage is not particularly frequent: less than 20 % of finite verbs are marked by one of these enclitics – around 15 % are marked by the dual, and 5 % by the paucal. They strongly tend to index human and higher order animate referents.

Usage of these number marking enclitics is not consistent either: if one is employed in reference to a particular group, the following reference to that group may or may not involve the enclitic in the verb. This is illustrated in the narrative sequence (42)–(43), where the group of horses is indexed in the first sentence by the paucal enclitic, but this is not used in the second sentence. Such sequences are not uncommon in the corpus.

(42) *warang-jirri   yilba/   ngarragi yawarda bala-loondi-rri/*
     sit-1EXC.NOM|CL forever 1SG.OBL horse   send-1SG.NOM|3PL.ACC|CL-PA
     'We stopped there for good, and I let the horses go.'

(43) *garndiwangoorroo bala-yirridi/*
     many   send-1EXC.NOM|3PL.ACC|CL
     'We let them all go.'

What motivates the use of these verbal number marking enclitics remains unclear. Certainly it does not seem to be motivated merely by considerations of reference tracking: in the case of (42), for instance, this can hardly be the motivation: the obligatory bound pronouns in the verb, as well as the discourse context, provides the information that the set of horses is plural (in fact, three). Nor is there any other set of horses in the discourse world that this set might be confused with. Counts over the first hundred or so instances of both the dual and the paucal enclitic in the textual corpus were made. This revealed that both number enclitics tended not to be associated with overt NPs in the same clause. This tendency was strongest for the dual marker – 75 % of instances did not involve a syntagmatically related NP; for the paucal marker only slightly over a half (51 %) did not involve an NP in the same clause. Interestingly, when an NP did occur in the clause indexed by a number enclitic there was a very strong tendency for it to be number marked by one or more of the strategies discussed in §2 and §3.

The two regular means of number marking in NPs are by means of free numerals and by number marking enclitics, both of which are strongly associated with human NPs and NPs serving in participant roles. Numerals occur in two main discourse contexts.

– In the specification of number of newly introduced entities (that typically fill participant roles in the clause they are introduced in), where the numeral behaves in the manner of an indefinite determiner (see note 1). This is illustrated by *garndiwangoorroo* 'many' in the final clause of (44) and the first clause of (45).

- In attribution of the number of an already introduced entity the numerosity of which is known or predictable (e.g. if reference is made to the denizens of a cattle property), thus typically specifying that the entirety of set was involved, as shown by the second clause of (45).

(44) *ward-jirri-yi/*   *biliga/*   *midim/*   *jalba-yirrra/*
go-1EXC.NOM|CL-DL middle meet   meet:up-1EXC.NOM|3PL.ACC|CL
garndiwangoorroo/ booloomani ward-birra   *ngilmangi/*
many   bullock   go-3PL.NOM|3SG.ACC|CL from:east
'We two went, and on the way met up with a group who were droving cattle west.'

(45) *niyaji-ya garndiwangoorroo/ warang-birri*   *yoowooloo/*
this-LOC many   sit-3PL.NOM|CL man
doow-yirrra/   *garndiwangoorroo-nyali*
get-1EXC.NOM|3PL.ACC|CL many-REP
doow-yirrra-rri/
get-1EXC.NOM|3PL.ACC|CL-PA
'There were lots of people camping there, and we got them all.'

Numerals are not normally used to distinguish and/or track referent sets; they are, that is, not (normally) used for distinguishing purposes.

The number marking NP enclitics are not very frequent in usage, and far less frequent than VP number enclitics, and somewhat less frequent than numerals greater than *yoowarni* 'one'. The dual NP enclitic occurs 101 times in the textual corpus referred to in note 8, making it just slightly less frequent than *garndiwirri* 'two' (some 130 attestations); the plural NP enclitic occurs 48 times, making it far less frequent than the numerals above *garndiwirri* 'two' (some two hundred instantiations). Put differently, number marking enclitics occur about once per ten participant NPs. (These are approximate figures, which include only verbal clauses.)

The two number marking enclitics show rather different patterns of usage in the textual corpus. Both are more often found on lexical than pronominal NPs, though for the dual about 20 % of instances occur on pronominal NPs, while for the plural the figure is about 2 %. And while almost 80 % of lexical NPs marked by the plural include a determiner, less than a half of lexical NPs marked by the dual do. Almost one third of the instances of the dual enclitic occur in NPs that also include the numeral *garndiwirri* 'two'. Surprisingly, just 6 (about an eighth) of the plural marked NPs include a numeral, and this is always *ngarloodoo* 'three', never a word for 'many'. My guess is that the collective sense coded by *-yarndi* PL accounts for the rarity of plural marked NPs with *garndiwa* 'many': as numerosity increases the likelihood of unity would seem to decrease.

In the narrative corpus the NP number marking enclitics are virtually restricted to phrases with specific and definite referents. Unlike numerals they are rarely found

on NPs introducing new entities: they almost always occur on given, previously introduced NPs. Examples (19), (22) and (46) are typical: the referents of the number-marked NPs are previously mentioned in the text. (47) is perhaps telling: on uttering *yoowooloo-yarndi-ga* 'man-PL-ERG' the speaker immediately corrects himself to say *yoowooloo ngarloodoo-ngga* 'man three-ERG': the men in question had been mentioned in the previous text (along with their having dug up the body), but are new to the addressee of the reported utterance (the narrator) at this point.

(46) *niyi-yarndi yawarda yilba-nyali   barn-jirrra-rri/*
     that-PL   horse   forever-REP return-1EXC.NOM|3PL.ACC|CL-PA
     gooroongalja-yirra/
     Brooking:Springs-ALL
     'We sent the horses back to Brooking Springs for good.'

(47) *ward-ji::/*   *miga-mi-ngarra/*   *yoowooloo-yarndi-ga/ yoowooloo*
     go-3SG.NOM|CL say-3SG.NOM|CL-1SG.OBL man-PL-ERG   man
     ngarloodoo-ngga thoorrbood-birrani/   *gijali/ andrew/ aa*
     three-ERG   dig:up-3PL.NOM|3SG.ACC|CL dead   Andrew um
     jaajin-ga   miga-mi-ngarra/
     sergeant-ERG say-3SG.NOM|CL-1SG.OBL
     'He came and told me that three men had dug up the dead body, Andrews,
     the sergeant, told me.'

There are just a few exceptions, where the referent is newly introduced to the narrative. For instance, example (10), the first sentence of a narrative, introduces the protagonists, the two young men. This is done in the unusual fashion of introducing them as though they are established identifiable entities through the use of the endophoric determiner *niyaji* 'this'. (Presumably this establishes a sort of immediacy as in the comparable phenomenon in English, further underlined by the use of *yoowayi* 'yes'.) However, in the majority of apparent exceptions the NP marked by the number enclitic is a part of the specification of the composition of a group that has already been broadly specified by another nominal expression in the same clause. (48), for example, begins a new episode in the narrative with the establishment of a new character set with the pronoun *ngidi* 1EXC.CRD in initial position; the previous episode recounted the exploits of a cattle rustler. Significantly, the number marking enclitics are employed on subsequent NPs that specify the composition of the first person group.

(48) *e/ ngidi/   ngidi/   bagimab-jidi-yi/*
     um 1EXC.CRD 1EXC.CRD pack:bags-1EXC.NOM|3SGACC|CL-DL
     policeman-joorroo-ngga/ ngarloodoo ngidi-yarndi   policeman/
     policeman-DU-ERG   three   1EXC.CRD-PL policeman

yoowooloo-yoorroo garndiwirri/ ngidi/   and/ policeman yoowarni/
Aborigine-DU   two   1EXC.CRD and policeman one
'We packed up our things, the three of us, two Aborigines and one
policeman.'

Similar remarks apply for (49), which introduces the man's wife into the text for the
first time. Initial reference to her is by the afterthought *goornboo joorloo* 'the woman
also'; only then does the speaker employ the dual enclitic in the coordinate NP
specifying the husband-wife dyad.

(49) *maningga doow-la/*   *ward-la:*
     night   get-1SG.NOM|3SG.ACC|CL go-1SG.NOM|3SG.ACC|CL
    mird-limi/   *goornboo joorloo/ goornboo-yoorroo:: yoowooloo/*
    tie-1SG.NOM|3SG.ACC|CL woman   also   woman-DU   man
    mird-loonmi-yi/   *girili-ya/ jayin-ngarri-ngga/*
    tie-1SG.NOM|3PL.ACC|CL-DL stick-LOC chain-COM-ERG
    'That night I got him, took him and tied him up, together with his wife, to a
    tree, with a chain.'

Numerals and number marking enclitics are thus used in similar environments;
both are used in (48), for instance. However, they are used in somewhat different
ways, with different meanings. In reference tracking, numerals are used to convey
the notion that the entirety of a set of the given cardinality is involved, whereas
number enclitics underline unity of the referent set. In introductive usage numerals
indicate plain cardinality of a set. NP number marking enclitics are not normally
used in introducing new entities into narratives; where they are so used, they indi-
cate the internal composition of the referent set already specified in more general
terms by a previous NP.

5 Conclusions

Number marking in Gooniyandi has by and large phrasal loci. The vast majority of
lexemes in the language are unspecified for number in the sense that they can be
used in reference to one or more instances of a particular category without any
morphological modification, inflectional or derivational. Lexical nominals in Gooni-
yandi, as in the typical Australian language, are transnumeral; they are set nouns in
the sense of Rijkhoff (2002: Chapter 2, 2003: 219–220, 2008). Verbs are also typically
unspecified for number. The major exceptions are derived stems, some of which
show inherent plurality, and certain grammatical nominals, in particular personal
pronouns.

Number marking of a given referent set can occur in various places in a clause. However, the marking directly concerns the referent set, and is not mediated through agreement of categories marked in the lexical items. Moreover, marking of number in one locus does not imply marking in another. Number marking at most loci is optional, the main exceptions being in pronouns, free and bound.

Number marking of types and sets of sets are poorly understood in Gooniyandi (and other Australian languages), and warrant further research. A range of other areas would benefit from further research. These include: the discourse usages of number marking enclitics on NPs (only partly understood) and enclitics on finite verbs;[10] apparent conflicts in number marking; other phenomena that are more or less related semantically or pragmatically to number (e.g. usage of *jooloo* 'also', *-nyooloo* ETC and *-ngarraya* TOO); and number as a cognitive category and its connection with number marking as a linguistic phenomenon. The latter is a particularly significant issue given the typical restricted numeral systems of Australian languages and widespread Whorfian interpretations concerning the non-salience of numbers to Aboriginal styles of cognition. It emerges clearly from this paper that despite having a restricted numeral system, and the inexactitude in reference of numerals, number is highly salient in Gooniyandi, and scarcely a clause goes by that is bereft of number indication.

Acknowledgements

Abbreviations and conventions

The following abbreviations and conventions are used in Gooniyandi example sentences:

ABL	ablative
ACC	accusative
CL	(verbal) classifier
COM	comitative

---

10 These number marking morphemes are not optional in the sense of (McGregor 2013b), and their usage or non-usage does not seem to be motivated by considerations of attention.

CRD	cardinal (a pronominal case form see § 2.2)
DAT	dative
DL	dual verbal enclitic
DU	dual postposition
DY	kinship dyad
ERG	ergative
ETC	etcetera
EXC	exclusive
FCT	factive mood
FUT	future
ICT	inceptive
INC	inclusive
IT	iterative
LOC	locative
NOM	nominative
NP	noun phrase
OBL	oblique
PA	paucal
PL	plural
POT	potential mood
PRG	progressive
PRS	present
REP	repeated
SG	singular
SUB	subjunctive
TOO	too, including, also
UNKN	unknown

The first three numerals indicate the three person categories. Morpheme boundaries in the classifier complex of Gooniyandi (see (2) in the paper) are indicated by |.

References

Allan, Keith. 1977. Classifiers. *Language* 53. 284–310.

Anon. 2004. *Yarrangi Gooniyandi thangarndi: Gooniyandi-English dictionary*. Unpublished manuscript. Fitzroy Crossing: Kimberley Language Resource Centre.

Bowern, Claire & Jason Zentz. 2012. Diversity in the numeral systems of Australian languages. *Anthropological Linguistics* 54 (2). 133–160.

Capell, Arthur. 1940. The classification of languages in north and north-west Australia. *Oceania* 10. 241–272, 404–433.

Dixon, Robert M. W. 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton.

Dixon, Robert M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.

Evans, Nicholas D. 2006. Dyad constructions. In Keith Brown (ed.), *Encyclopaedia of language and linguistics*. Volume 4, 24–27. Oxford: Elsevier.

Fraser, John. 1890. Some remarks on the Australian languages. *Royal Society of New South Wales, Journal and Proceedings* 24. 231–253.

Hale, Kenneth. 1975. Gaps in grammar and culture. In M. Dale Kinkade, Kenneth Hale & Oswald Werner (eds.), *Linguistics and anthropology: in honor of C. F. Voegelin*, 295–315. Lisse: The Peter de Ridder Press.

Hopkins, Elizabeth B. 1986. Pronouns and pronoun fusions in Yaouré. In Ursula Wiesemann (ed.), *Pronominal systems*, 191–203. Tübingen: Gunter Narr.

McGregor, William B. 1988. Mood and subordination in Kuniyanti. In Peter Austin (ed.), *Complex sentence constructions in Australian languages*, 37–67. Amsterdam: John Benjamins.

McGregor, William B. 1989a. Gooniyandi mother-in-law "language": dialect, register or code? In U. Ammon (ed.), *Status and function of languages and language varieties*, 630–656. Berlin & New York: Mouton de Gruyter.

McGregor, William B. 1989b. Phrase fracturing in Gooniyandi. In L Marácz & Peter Muysken (eds.), *Configurationality: the typology of asymmetries*, 207–222. Dordrecht: Foris Publications.

McGregor, William B. 1990a. *A functional grammar of Gooniyandi*. Amsterdam: John Benjamins.

McGregor, William B. 1990b. An impersonal construction in Gooniyandi, Western Australia. *Word* 41 (2). 161–184.

McGregor, William B. 1992. The noun phrase as a grammatical category in (some) Australian languages: a reply to Mark Harvey. *Australian Journal of Linguistics* 12. 315–319.

McGregor, William B. 1996a. Dyadic and polyadic kin terms in Gooniyandi. *Anthropological Linguistics* 38. 216–247.

McGregor, William B. 1996b. The pronominal system of Gooniyandi and Bunuba. In William B. McGregor (ed.), *Studies in Kimberley languages in honour of Howard Coate*, 159–173. München & Newcastle: Lincom Europa.

McGregor, William B. 1997. Functions of noun phrase discontinuity in Gooniyandi. *Functions of Language* 4 (1). 83–114.

McGregor, William B. 2002. *Verb classification in Australian languages*. Berlin and New York: Mouton de Gruyter.

McGregor, William B. 2003. The nothing that is, the zero that isn't. *Studia Linguistica* 57 (2). 75–119.

McGregor, William B. 2004. Numerals in Australian languages. Presented at *Workshop on Numerals in the World's Languages*, Max Planck Institute for Evolutionary Anthropology, Leipzig.

McGregor, William B. 2005. Quantifying depictive secondary predicates in Australian languages. In Nikolaus P. Himmelmann & Eva Schultze-Berndt (eds.), *Secondary predication and adverbial modification: the typology of depictives*, 173–200. Oxford: Oxford University Press.

McGregor, William B. 2007. Number words and number symbols in Australian languages. Paper presented at Centre de Recherches Linguistiques sur l'Asie Orientale, EHESS-CNRS, Paris.

McGregor, William B. 2013a. Lexical categories in Gooniyandi, Kimberley, Western Australia. In Jan Rijkhoff & Eva van Lier (eds.), *Flexible word classes: typological studies of underspecified parts of speech*, 221–246. Oxford and New York: Oxford University Press.

McGregor, William B. 2013b. Optionality in grammar and language use. *Linguistics* 51 (6). 1147–1204.

Merlan, Francesca & Jeffrey Heath. 1982. Dyadic kinship terms. In Jeffrey Heath, Francesca Merlan & Alan Rumsey (eds.), *Languages of kinship in Aboriginal Australia*, 107–124. Sydney: Univeristy of Sydney.

Nekes, Hermann & Ernest A. Worms. 2006. *Australian languages*. Edited by William B. McGregor. Berlin and New York: Mouton de Gruyter.

O'Grady, Geoffrey N., Carl F. Voegelin & Frances M. Voegelin. 1966. Languages of the world: Indo-Pacific fascicle 6. *Anthropological Linguistics* 8 (ii). 1–197.

Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.

Rijkhoff, Jan. 2003. When can a language have nouns and verbs? *Acta Linguistica Hafniensia* 35. 7–38.

Rijkhoff, Jan. 2008. On flexible and rigid nouns. *Studies in Language* 32 (3). 727–752.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Wurm, Stephen A. 1972. *Languages of Australia and Tasmania*. The Hague: Mouton.

David Gil

13 Number in Indonesian

Abstract: This chapter presents a survey of number in four colloquial varieties of Indonesian: Riau Indonesian, Jakarta Indonesian, Kupang Malay and Papuan Malay. In general, number plays a very minor role in the basic grammatical organization of the language. Nevertheless, a wide range of strategies are available for the optional expression of number, with a substantial amount of cross-dialectal variation. Some of the number-related constructions discussed in this chapter include reduplication, numeral classifiers, a plural word, associative plurals, inclusory pronouns, dual pronouns and argument indexation.

1 Overview

This chapter presents a survey of number in Indonesian, as exemplified by 4 geographically-delimited colloquial varieties: Riau Indonesian, Jakarta Indonesian, Kupang Malay and Papuan Malay.[1]

1.1 The Indonesian linguistic landscape

Indonesian, together with its alter ego Malay, is one of the world's major languages, spoken by well over 200 million people, primarily in Indonesia, Malaysia, Singapore and Brunei, plus also by smaller groups in other countries. Within Indonesia, the Indonesian language is spoken alongside nearly 800 other languages. Indonesian itself exists in myriad varieties, and speakers are typically diglossic or even multiglossic, alternating between Standard Indonesian and one or more colloquial varieties of Malay/Indonesian. These latter colloquial varieties belong to two main sociolinguistic categories: traditional dialects spoken by ethnic Malays primarily on the islands of Sumatra and Borneo, and koiné contact varieties spoken by many or most Indonesians throughout the archipelago. This chapter focuses on 4 exemplars of the latter category, contact varieties of Indonesian.[2]

---

1 The terms Indonesian and Malay are commonly used to refer to different varieties of what is a single macro-language, Malay/Indonesian. The choice between these two terms reflects established traditions and is motivated by a variety of largely sociolinguistic factors; it does not correlate with any purely linguistic features, nor does it reflect a historical split of Malay/Indonesian into two branches.
2 The focus on colloquial varieties of Indonesian to the exclusion of the standard language is justified by their ontological primacy, as reflected in their earlier acquisition, more widespread distribution, and greater naturalness. Colloquial varieties of Indonesian exhibit a variety of interest-

Proceeding from west to east, Riau Indonesian is spoken by a few million people in east central Sumatra and some smaller adjacent islands, in the provinces of Riau and Kepulauan Riau; some previous studies include Gil (1994, 2002, 2004, 2005a,b, 2009, 2013b, 2020). Jakarta Indonesian is spoken by up to 20 million people in the capital city Jakarta and its environs, and also by many mostly younger Indonesians throughout the country as a somewhat trendy lingua franca, as a result of which it is occasionally, somewhat misleadingly, referred to simply as "colloquial Indonesian"; some descriptions include a reference grammar by Sneddon (2006), plus also Wouk (1989, 1999), Tjung (2006), Hidayat (2010), and Conners, Bowden and Gil (2015). Kupang Malay is spoken by one million or more people in the city of Kupang and some neighboring parts of the western half of the island of Timor in Nusa Tenggara Timur province, and to a lesser extent also as an alternative to other local varieties of Indonesian in other parts of the province; some previous studies include Steinhauer (1983), Mboek (1984), and Jacob and Grimes (2011). Finally, Papuan Malay is spoken on the western half of the island of New Guinea, in what is now sometimes referred to as Tanah Papua, or the Land of Papua, currently encompassing the provinces of Papua and Papua Barat; some descriptions include a reference grammar by Kluge (2017) and also Donohue and Sawaki (2007) and Saragih (2012).

The varieties of Indonesian discussed in this chapter vary greatly in mutual intelligibility, depending on particular speakers and circumstances. Of course, they share many features, so for ease of exposition, in those cases where a statement applies to all four varieties under consideration, we shall simply use the term "Indonesian" (keeping in mind that there may well be other varieties to which the statement in question may not apply). However, as we shall see below, these four varieties also exhibit a substantial amount of variation with respect to the features under consideration. The descriptions presented in this paper draw largely from the author's familiarity with these four varieties of Indonesian, based mostly on immersion, but supplemented also with specific targeted elicitation. Most of the data cited in this paper are from naturalistic corpora: Gil, Tadmor, Bowden and Taylor (2015) for Jakarta Indonesian and Kupang Malay, and the author's own diary studies, recordings, and social-media corpora, mostly for Riau Indonesian and Papuan Malay.[3]

---

ing and less familiar grammatical features in the domain of number that are not present in Standard Indonesian.

3 The naturalistic data is presented here in four lines: (i) a transcription, making use of a conventionalized orthography; (ii) an interlinear gloss; (iii) a free translation; and (iv) a description of the context in which the utterance occurred, enclosed in square brackets. In one case, example (3), involving a ludling, the transcription represents the utterance providing the input to the ludling, while an additional line above it displays the utterance as it actually occurred, reflecting the output of the ludling. In several other cases, involving written texts from social media, the transcription represents the utterance in a normalized orthography, while an additional line above it shows the utterance as it was actually typed, with variant spellings, punctuation and capitalization.

1.2 A typological profile of Indonesian

In accordance with Greenberg's (1963) word-order typology, Standard Malay emerges as a rather typical head-initial language, falling into his well-populated #9 word-order category, alongside European languages such as French and Greek, with SVO, PrepN, NG and NA basic word orders. The varieties of Indonesian under consideration here are broadly similar, though Kupang and Papuan Malay have basic GN rather than NG order. However, as argued in Conners, Bowden and Gil (2015) with reference to the Jakarta dialect, Indonesian is a "wolf in sheep's clothing", in which a superficial typological similarity to Standard Average European masks profound differences in grammatical organization. Accordingly, Gil (2017), with reference to Riau Indonesian, proposes a characterization of Indonesian as coming close to instantiating the prototype of IMA language, in which the three capital letters stand for Isolating, Monocategorial and Associational respectively.

The relatively IMA nature of Indonesian may be illustrated in a simple sentence such as the following:

(1) *Ayam   makan*
     chicken eat

The isolating nature of Indonesian is evident in the fact that, as exemplified in (1), sentences typically consist mostly or entirely of free-standing content words, without any additional bound morphemes; as argued in Gil (2020), Riau Indonesian, and by extension other varieties of Indonesian, come as close as any other language to instantiating pure isolating structure. The monocategorial character of Indonesian, argued for in Gil (1994, 2005b, 2013b) for Riau Indonesian, results from the absence of any distinction between lexical categories such as noun, verb, adjective and adposition, or between lexical categories and their would-be phrasal projections; thus, in (1) above, *ayam* 'chicken' and *makan* 'eat' exhibit identical grammatical behaviour, both in (1) itself and in other expressions in which they occur. Finally, the associational nature of Indonesian, argued for in Gil (1994, 2005b, 2017) for Riau Indonesian, reflects the fact that when words or longer expressions are brought together, the meaning of the resulting combination is semantically underspecified, and can be anything that has to do, in one way or another, with the meanings of the constituent parts; for example, in (1) the chicken may be either the agent or the patient of the eating, and the expression as a whole may refer to either an activity, e.g. 'The chicken is eating', or a thing, e.g. 'The chicken that is eating'.

1.3 Number in Indonesian

The typological profile of Indonesian as a relatively IMA language has far-reaching implications vis à vis the category of number. In a nutshell, number plays a minor,

indeed a vanishingly small role in the grammar of Indonesian. As an isolating language, there is simply little or no room for a grammatical category whose primary manifestation, by definition, involves bound morphology.

The paucity of grammatical number in Indonesian is clearly evident in example (1) above. In (1), *ayam* may be understood as either count or mass, and, when count, as either singular or plural; similarly, *makan*, with no tense or aspect marking present, may be interpreted as denoting either an undifferentiated generic activity, or, alternatively, a single punctual activity or a series of iterated ones. Moreover, the form of *makan* says nothing about whether *ayam* is mass, singular or plural; there is no structural relationship between *ayam* and *makan* of the kind that might be considered to involve number agreement.

So what is there, then, in Indonesian, that might be subsumed under the notion of number? Essentially, a potpourri of phenomena, some lexical, others grammatical or discourse-oriented, many of which differ from one variety of Indonesian to another, all of which are peripheral to the fundamental ground plans of the language. Pronouns are the only domain in which obligatory number marking can be observed, but even there it is mostly of limited scope, except in Papuan Malay, which has also innovated a series of dual pronouns (Section 2.2). Numeral classifiers are present but optional (Section 2.3.1). Certain words may be associated with inherent number, but there is more fluidity here than in many other languages (Section 2.3.2). A particularly rich source of cross-dialectal variation is provided by several different constructions which may be characterized as inclusory pronouns and/or associative plurals; some of these constructions have taken on additional functions, such as the Kupang Malay plural word *dong* (Section 2.3.3). In Jakarta Indonesian, there is a plural word *pada* whose function is to mark its host as having a plural argument (Section 2.4.1). Reduplication, illustrated here from Riau Indonesian, has a wide range of functions, many of which can be reduced to a basic meaning involving plurality and distributivity (Section 2.4.2). Finally, in Papuan Malay, pronouns have developed the additional role of argument indices, and as such, serve to mark the number of their host expressions' arguments (Section 3). Although Indonesian is supposedly a well-known language, much of the above material constitutes novel descriptions of hitherto unreported facts.

The organization of this chapter and the descriptions therein make reference to etic categories, or comparative concepts, in the sense of Haspelmath (2010); this perspective is adopted in order to render the description of Indonesian easily accessible to typologists, and comparable to the other chapters in this volume. It should be kept in mind, however, that this comparative typological approach often ends up doing violence to the way the language actually is; that is to say, it comes into conflict with the emic, or language-specific descriptive categories of Indonesian. Some of the conflicts between etic and emic approaches are touched upon in this chapter. Thus, for Papuan Malay the etic categories of inclusory pronoun and associative plural are argued to be instantiated by one and the same emic construction (Sec-

tion 2.3.3), while for Riau Indonesian, it is argued that a more insightful language-specific description of reduplication would treat it as a single construction applying across the comparative concepts of nominal and verbal number (Section 2.4.2).

2 Pronominal, nominal and verbal number

Due to the isolating nature of Indonesian, there is no clear-cut distinction between morphology and syntax, and hence no easily-drawable distinction between number, a primarily morphological notion, and quantification, manifest primarily in the lexicon and syntax. Accordingly, this section brings together phenomena that, in other languages, are more appropriately afforded separate treatments of a morphological and syntactic nature respectively.

2.1 Generalities

Because of its monocategorial nature, Indonesian does not clearly distinguish between pronouns, nouns and verbs. For the most part they exhibit similar grammatical behaviour, both generally and with respect to number; whatever differences that are observable can be accounted for with reference to purely semantic features. Accordingly, in this chapter, we shall use terms such as pronoun, noun and verb, to refer to semantic categories, without prejudice with regard to whether such categories also correspond to morphosyntactic ones.

2.2 Pronominal number

The class of pronouns in Indonesian may be defined semantically as the class of words that bear specification for the value of the feature of person. Most words in Indonesian, e.g. *ayam* and *makan* in (1), usually have 3rd person reference, but may in appropriate contexts also assume 2nd or 1st person reference. In contrast, pronouns are specified for person and may not generally assume person values other than those they are specified for.

For three of the four varieties of Indonesian described here, Riau Indonesian, Jakarta Indonesian and Kupang Malay, the class of pronouns is noteworthy for the following two additional properties, which are generally characteristic of Mainland Southeast Asian languages as well as other languages of Indonesia. First, it is a semi-open class, in that new forms are frequently being introduced and dropping out of use; such variability seems to be motivated largely by considerations of politeness and honorificity. The second related property is that the use of pronouns is never obligatory, and in fact is rather infrequent; this is the feature characterized

by Helmbrecht (2005) as "pronoun avoidance". In contrast, with respect to these two properties, Papuan Malay bears a closer resemblance to Standard Average European languages than it does to other varieties of Indonesian.

The most common pronouns in each of the four varieties of Indonesian are presented in Tables 1–4 below:

Tab. 1: Pronouns in Riau Indonesian.

	singular	plural	
		inclusive	exclusive
1st person	aku, saya	kita	kami
2nd person	kau, engkau	kalian	
3rd person	dia		

Tab. 2: Pronouns in Jakarta Indonesian.

	singular	plural
1st person	gua, aku, saya, ogut	kita
2nd person	lu, kamu	
3rd person	dia	

Tab. 3: Pronouns in Kupang Malay.

	singular	plural	
		inclusive	exclusive
1st person	beta [be]	kotong	botong
2nd person	lu	bosong	
3rd person	dia [de]	dorang [dong]	

Tab. 4: Pronouns in Papuan Malay.

	singular	plural	dual
1st person	saya [sa]	torang [tong]	tendua
2nd person	koi [ko]	kamu [kam]	kamdua
3rd person	dia [de]	dorang [dong]	dendua
3rd person inanimate	akan		

Perhaps the most noticeable fact about the above tables is how much the pronominal systems differ across the four otherwise relatively closely-related dialects, with respect to both matter and pattern.[4]

In general, pronouns are the only domain in Indonesian where there is obligatory number specification. However, even here, such specification is contingent on person and animacy, in accordance with the familiar referential hierarchy, see Silverstein (1976) and many others:

(2) *The referential hierarchy and obligatory number specification*

pronoun: 1st person	pronoun: 2nd person	pronoun: 3rd person animate	pronoun: 3rd person inanimate	other
Jakarta Indonesian	Riau Indonesian	Papuan Malay	Kupang Malay	

In accordance with (2), each dialect of Indonesian obligatorily specifies number for positions higher, or to the left, of its indicated cut-off point. Thus, number is specified for 1st person only in Jakarta Indonesian, for 1st and 2nd person only in Riau Indonesian, for 1st, 2nd and 3rd person animate only in Papuan Malay, and for all pronouns but no other words in Kupang Malay.

A further difference between the four dialects pertains to the number of values of the number feature. Whereas most varieties of Indonesian distinguish just between singular and plural pronouns, Papuan Malay also has a series of dual pronouns.[5] These forms are transparently derived from a periphrastic construction involving the

---

4 Given the fluid nature of Indonesian pronouns as a semi-open word class, Tables 1–4 can only present a simplified picture of a state of affairs that is actually considerably more complex along a number of dimensions. One major analytical challenge is posed by the distinction between "basic" meanings of pronouns, represented in Tables 1–4, and various arguably secondary or derived meanings that such pronouns may also assume in certain contexts. For example, in Riau Indonesian, the 1st person plural exclusive *kami* can also be used as a polite, self-deprecating 1st person singular. Elsewhere, in Papuan Malay, the 2nd person plural *kamu* can also be used as a more polite variant of the 2nd person singular (reminiscent of the well-known "T/V alternation" in many European languages), though it is not obvious whether this usage should be considered as internal to Papuan Malay or alternatively as instantiating code-switching between Papuan Malay, with *ko*, and some other language variety such as Jakarta Indonesian or Standard Indonesian, in which *kamu* may also be understood as singular.

5 Kluge (2017) contains no mention of the dual pronouns described here. A possible explanation for this might involve regional variation within Papuan Malay; whereas Kluge's description is based on the variety spoken in Sarmi, Papua province, the variety that I am most familiar with is that spoken in Manokwari, Papua Barat province. It is likely that the presence of the dual pronouns in

juxtaposition of a plural pronoun with a numeral, for example *tong dua* 'we two', *tong tiga* 'we three', *tong empat* 'we four', and so forth. However, in the case of 'two', but no numerals higher than 'two', the periphrastic forms *tong dua* 'we two' and *dong dua* 'they two' are optionally reduced to *tendua* (1DU) and *dendua* (3DU).

Phonological evidence for the reduction is provided by the replacement of [o] with the neutral vowel [e], and the assimilation of [ŋ] to [n] before [d]. In addition, two sources of evidence, both of a paralinguistic nature, support the analysis of *tendua and dendua* as constituting single words. The first is the Bahasose ludling (named after how it applies to the word *bahasa* 'language'), described in Gil (2013a). The Bahasose ludling applies to each word as a whole, replacing the final disyllabic foot $-C_1V_1C_2V_2C_3$ with the sequence $-C_1V_1C_2oC_2eC_3$ (where C stands for zero, one or two consonants and V for zero or one vowels), as in the following naturalistic example:[6]

(3) *Langsosen tendu[w]o[w]e putotes*
     Langsong tendua putus
     straight 1DU   break.off
     '... and we immediately broke up.'
     [Speaker describing his love life.]

As shown above, the dual pronoun *tendua* is treated by the ludling as a single word: if the form to which the ludling applied were the periphrastic *tong dua*, the expected output would have also consisted of two words, *tongonge duwowe*. The second source of evidence is from naturalistic orthography: in the absence of prescriptive norms, people often choose to write the forms in question as single words, as in the following example:

(4) *Pit bhaa empong tendua ada main wifi krim foto khaa*
     Pit ba, Empong tendua ada   main wifi, kirim foto   ka
     Pit IRRIT Empong 1DU   PROG play wifi send photo Q
     'Pit, Empong and I are online, please send some photos.'
     [Facebook chat: writer addressing me (using my nickname *Pit*).]

As evident in the above example, *tendua* is written as a single word. In this respect, it contrasts with its non-reduced periphrastic counterparts, written separately as in example (56) below. The reduced and coalesced nature of *tendua* and *dendua* suggests that these are indeed bona fide dual forms rather than periphrastic pronoun-plus-numeral constructions. (As for the 2nd person form *kamdua*, there is no phono-

---

the Manokwari subdialect, at least, is a recent development resulting from contact with local languages that have more widespread dual number, such as, for example, Biak.

6 In the first line of example (3), representing the actual pronunciation of the ludling, the two semi-vowels, written as *[w]*, are inserted automatically in order to separate the adjacent vowels.

logical evidence for reduction; in the absence of corpus data from ludlings and naturalistic orthography, its analysis as a dual is by analogy to the 1st and 3rd person forms.) As in many other languages with duals, cf. Corbett (2000:43–50), the use of the dual in Papuan Malay is optional, and a plural form can always be used in its place. Impressionistically, though, the use of the dual is quite frequent in contexts where precisely two persons are involved.

The forms assumed by the pronouns in Tables 1–4 also exhibit some interesting patterns. In the eastern varieties, Kupang and Papuan Malay, some of the pronouns occur in two variants, a "long" disyllabic variant and a "short" monosyllabic variant, the latter enclosed in Tables 3 and 4 in square brackets. As far as I have been able to ascertain, there are no syntactic or semantic factors conditioning the choice between these two forms. Also in the eastern varieties, many of the plural pronouns end in *-ang* or *-ong*. These latter forms are historically relatively recent reductions of a periphrastic construction of the form *PRON orang*, where *PRON* is a pronoun and *orang* the word for 'person'; in other varieties of Malay and Indonesian, the transparent periphrastic nature of the construction is still in evidence, for example the Sabah Malay 3rd person plural pronoun *dia orang*, consisting of 3rd person *dia* plus *orang* 'person'. Thus, historically if not actually synchronically, *orang* functions as a formative for the creation of plural pronouns. A similar construction is also evident in Hokkien Chinese with *lâng* 'person' and in Tok Pisin *pela* from English *fellow*, suggesting that its origin may be in a Southeast Asian trade language. Synchronically, however, it is not clear why the addition of *orang*, itself unspecified for number, should have a pluralizing effect.[7]

2.3 Nominal number

Nouns in Indonesian are generally unspecified for number; as noted above an expression such as *ayam* may be understood as mass, singular or plural.[8] However,

---

[7] Compounding the puzzle is an additional fact regarding the combination of *orang* with the pronoun *kita*, whose most common meaning is 1st person plural. While in Kupang and Papuan Malay, the collocation *kita orang* forms the basis for the 1st person plural pronouns *kotong* and *torang* respectively, in at least two other varieties, Southern Lampung Indonesian and Kuala Lumpur Malay, adding *orang* to *kita* has a different effect: in addition to effecting a plural interpretation, it forces an exclusive interpretation: in both these dialects, *kita* (1PL.INCL) contrasts with *kitorang* (1PL.EXCL). Curiously, a somewhat similar pattern is once again evident also in Tok Pisin with *mipela* (1PL.EXCL). Again, I have no explanation for these facts.

[8] Looking beyond the four varieties of Indonesian that form the focus of this chapter, a partial exception to the underspecification of nominal number may be observed in Singaporean Malay and perhaps other varieties, in the form of a *definiteness effect*. Specifically, while *ayam* on its own is unspecified for number, when a definite determiner is added, e.g. *ayam itu* (chicken DEM:DEM.DIST), the resulting expression is understood as either mass or singular, but not as plural. In order to force a plural interpretation 'those chickens', the noun must be reduplicated, *ayam-ayam itu* – see Gil

Indonesian does possess several morphosyntactic strategies whose meanings are connected in one way or another with the notion of nominal number.

2.3.1 Numeral classifiers

When a noun is modified by a numeral, a numeral classifier may optionally be present; some examples include *dua (ekor) ayam* (two CLF chicken) 'two chickens', *tiga (orang) guru* (three CLF teacher) 'three teachers', *empat (botol) air* (four bottle water) 'four bottles of water'.

While mensural classifiers resemble those in many other languages, sortal classifiers play a smaller role than in most Mainland Southeast Asian languages: they are few in number, their use is relatively infrequent, and they occur in a more limited range of syntactic contexts, specifically, only with numerals. Unlike in many Mainland Southeast Asian languages, they do not occur with other quantifiers, demonstratives, or other attributive modifiers, and they do not occur "bare", that is to say, without any modifier whatsoever, as singulative markers on the noun, as, for example, in Vietnamese *con chó* (CLF dog) 'the dog'. The presence of the quantified noun is not obligatory; for example, the expression *dua ekor* (two CLF) may occur on its own as a numeral-plus-classifier construction meaning 'two (animals)'.

Both sortal and mensural classifiers do double service as regular content words; however, in the case of sortal classifiers, their semantics changes. For example, in *dua ekor ayam*, the classifier *ekor*, used for animals, is also an ordinary content word meaning 'tail'; accordingly, the expression *dua ekor ayam* could potentially also mean 'two chicken tails'. Two other common classifiers whose meanings change in a similar way are *biji*, for small objects, also 'seed', and *buah*, for larger objects, also 'fruit'.

Given the optionality of the classifier in the noun-plus-numeral construction, the question arises what factors govern its presence or absence in such environments. In an analysis of a literary Malay text, Hopper (1986) argues that the relevant factors are not grammatical or semantic but rather pragmatic; specifically, the presence of the classifier is associated with nouns of greater discourse persistence. Whether or not similar generalizations might be valid also for the varieties of colloquial Indonesian under consideration in this chapter must be left for future research.

2.3.2 Lexically inherent number

Some nouns are inherently associated with particular values of number. To begin with, as in probably all languages, nouns come with countability preferences, re-

---

(2003) for further discussion. However, such definiteness effects are not attested in any of the koiné varieties of Indonesian.

flecting the relative availability of count and mass interpretations, cf. Allan (1980), Gil (1996). For example, while *ayam* 'chicken' is mid-range on the countability scale, readily allowing for both count and mass interpretations, *guru* 'teacher' is of high countability, while *air* 'water' is of low countability. Compared to other languages, however, countability preferences bear relatively few grammatical consequences. For example, the use of numeral classifiers, exemplified in the previous section, is optional rather than obligatory for all nouns, regardless of their countability preference; thus, both high countability preference *guru* 'teacher' and low countability preference *air* 'water', when quantified with a numeral, can occur either with or without a numeral classifier.

Nevertheless, nouns of high and low countability preference differ in one important aspect of their morphosyntactic behaviour: whereas nouns of high countability preference may readily occur in construction with adjectives of size and shape, for example *guru tinggi* (teacher tall) 'tall teacher', *rumah bulat* (house round) 'round house', nouns of low countability preference cannot generally do so, for example *?*air tinggi* (water tall) 'tall water', *?*pasir bulat* (sand round) 'round sand'. Given that adjectives of size and shape involve built-in reference to natural units of individuation, the above contrast shows that nouns of high countability preference in Indonesian such as *guru* 'teacher' and *rumah* 'house' are indeed inherently associated with such natural countable units. This in turn entails that whatever the function of the (optional) numeral classifiers in Indonesian may be, it is not to impose a unit of countability on an erstwhile mass noun lacking such natural units. That is to say, even in the absence of a numeral classifier, nouns such as *guru* 'teacher' and *ayam* 'chicken', are inherently associated with natural units of counting; indeed, in an appropriate context, even nouns of low countability preference such as *air* 'water' may be associated with such units. As argued in Gil (1996), similar facts in other languages such as Vietnamese and Tagalog provide strong evidence against approaches that account for transnumerality, or the underspecification of number, by characterizing the nouns in question as having mass interpretations, as for example in Quine (1969:35–38), Link (1991) and Chierchia (2010).

Focusing now on the nouns of mid or high countability preference, as noted above, most such nouns may be interpreted equally readily as either singular or plural; however, there are a few exceptions to this generalization. First, as in other languages, there are nouns that denote a collective, for example *kelompok* 'group', *Persipura* (the name of a football team).[9] Perhaps more than in other more familiar languages, however, there is a certain degree of fluidity, with many nouns allowing

---

9 As shown in Gil (1996), the term 'collective' is used in a bewildering variety of ways to describe a wide range of linguistic phenomena. In this chapter, the term is used to refer to words referring to an entity consisting of a plurality of constituent parts where, crucially, the whole is more than just the some of the parts; this usage of the term is accordingly referred to in Gil (1996) as "nonadditive".

for both collective and individuated singular or plural interpretations. Such variation comes to the fore in loan words that were collective in the source language while allowing individuated interpretations in Indonesian. One such example is *tentara,* which was borrowed into Early Malay from Sanskrit *tantra* 'army' (Tadmor 2009); in Indonesian, however, it can mean either 'army' or 'soldier'. A modern counterpart of this is *polisi*, from Dutch *politie* 'police', which in Indonesian can mean either 'police' or 'policeman/policewoman'.

The origin of this fluidity would appear to lie in the availability of a zero-marking option for possessive constructions, including partitive ones. Consider the following example from the dialect of Indonesian spoken in Bandung, which is very similar to Jakarta Indonesian, illustrating the use of *kaum* 'community', borrowed from Arabic *qawm*, 'community' or 'nation':

(5) *Oh, kita kaum   wanita sih*
     oh 1PL community woman EMPH
     'Oh, because we are (of) the community of women.'
     [Speaker explaining why she and her friend are so noisy.]

Because the partitive relationship, indicated in parentheses in the translation above, is not overtly expressed, the collective noun *kaum* 'community' is prone to reanalysis as an ordinary noun, denoting, simply, 'person' or 'people'. Hence, also, expressions such as *banyak kaum* (many community) may be understood as either 'many communities' or 'many of the community'.

As in other languages, there are also expressions that are inherently singular, the most important instance of this being proper nouns denoting people, places, events and other name-bearing entities. If plural reference is required, such as to a set of people called *Rudi,* various ad hoc periphrastic devices are drawn upon. In the two western varieties, Riau and Jakarta Indonesian, there is also a formal device for creating proper nouns denoting people, namely *si*. Most commonly, *si* occurs in front of a proper noun, e.g. *si Rudi*, in which case it is redundant. However, in other cases, it occurs in front of other kinds of expressions in order to convert them into proper nouns denoting people. One instance of this involves the conventionalized personification of animals in folk takes, e.g. *si kancil* (PERS mousedeer) – cf. the English *Mr. Mousedeer*. Another instance involves the ad hoc creation of names based on attributes, e.g. *si ompong* (PERS snaggletooth); such attributes may also be phrasal, e.g. *si rambut kuning* (PERS hair yellow). Notably, the reference of such expressions is invariably singular. For example, *si ompong* can only refer to one person; it thus contrasts with the similar expression *yang ompong* (REL snaggletooth), formed with the reifier/relativizer *yang*, which can refer to either one or more than one person, 'the snaggletooth one(s)'.[10]

---

10 As argued in Gil (2020), in Riau Indonesian the personal marker *si* is more appropriately analyzed as a bound form, contrary to the standard orthography in which it is written separately.

A further case of a somewhat more complex nature is that of address terms; a class of expressions consisting mainly of kinship terms, but also various other expressions such as titles, names of professions, and terms of ethnicity; see Conners, Brugman and Adams (2016) and Adams and Conners (to appear) for some discussion of these. The defining characteristic of address terms is that, in addition to the usual 3rd person reference, they may also be used freely with 1st and 2nd person reference. On the one hand, their common use with 1st and 2nd reference distinguishes them from 3rd person pronouns, whose use in 1st and 2nd person contexts is impossible, as well as from most other words, whose use in 1st and 2nd person contexts is infrequent. On the other hand, their widespread use with 3rd person reference sets them apart from 1st and 2nd person pronouns, whose use in 3rd person contexts is impossible. Address terms would appear to occupy a middle ground in-between pronouns and other nouns, and indeed, they are often used in contexts where a Standard Average European language might use pronouns, or alternatively generic nouns with meanings such as 'man', 'woman', 'child' or so forth.

Address terms typically occur in two distinctive environments, vocative and preposed to proper nouns denoting individuals. As vocatives, they are generally unspecified for number. For example, in Papuan Malay, *kaka* (elder.sibling) can be used to call one or more persons of comparable or somewhat older age relative to the speaker. In the other three varieties of Indonesian, vocative address terms are often truncated to monosyllabic form. For example, in Riau Indonesian, *abang* (elder.brother) is shortened to *bang* in order to call one or more persons. However, preposed to proper nouns, with 1st, 2nd or 3rd person reference, the construction as a whole inherits its obligatory singular interpretation from the proper noun. While this construction is relatively infrequent in Papuan Malay, it is very common in the other three dialects. For example, in Riau Indonesian, a person named *Rudi* might be referred to as either *abang Rudi* or *bang Rudi*; in addition, the name itself might be truncated, resulting in the form *bang Rud*. As is the case for proper nouns, the resulting construction is specified as singular.

In addition to the above two distinctive environments, address terms also occur in other positions, just like any other words in the language. In such cases, though, the facts regarding number are rather more complex, exhibiting variation across dialects, speakers and lexical items. Whereas some kinship terms, such as those whose basic reference is to siblings, are generally unspecified for number, others, most commonly those whose basic reference is to parents, are often understood as inherently singular, even when used to refer to non-relatives. For example, when occurring in typical argument position, nouns such as Riau Indonesian *ibuk*, Jakarta

---

Indeed, in some cases, the combination of *si-*, or its phonologically reduced variant *i-*, plus a truncated monosyllabic form of a proper noun, may become lexicalized as a new proper noun, which of course also has invariable singular reference, for example *Paisal* (original name) > *Pai* (truncated variant) > *i-Pai* (personal marker added) > *Ipai* (reanalysis as new name).

Indonesian and Kupang Malay *ibu*, and Papuan Malay *mama*, all meaning 'mother', or simply 'woman', are usually interpreted as singular; if a plural interpretation is desired, the noun is more likely to be reduplicated. In such cases, it would seem as though the prototypically unique identity of 'mother' construed as the kinship term within the traditional nuclear family has been generalized also to the extended or "classificatory" usages of 'mother' referring simply to 'woman'.

A final and somewhat different case is provided by dyadic kinship terms. In all four Indonesian varieties under consideration there is a special term referring to two or more persons who are siblings. In Riau Indonesian, the term is formed by reduplication of the word for 'younger sibling' *adik* plus intercalation of the medial voice prefix *ber-*, yielding *adik-beradik*. In contrast, in the other three varieties, the term is formed by compounding the words for younger sibling and older sibling respectively, for example Papuan Malay *ade-kaka*.

2.3.3 Inclusory pronouns, associative plurals and their ilk

A prototypical inclusory pronoun construction is one such as is instantiated by the Russian *my s Ivanom* (1NOM.PL with John:INSTR.SG) 'me/us and John', where the pronoun *my* includes *Ivanom* in its reference; see Lichtenberk (2000), Haspelmath (2004) and others. A prototypical associative plural construction is one like the Hungarian *Pálék* (Paul:APL) 'Paul and associates', where the proper noun *Pál* is the focal referent of a set containing additional members associated with *Pál*; see Daniel (2000), Moravcsik (2003), and Daniel and Moravcsik (2005). Indonesian provides a rich source for both inclusory pronoun and associative plural constructions, with substantial cross-dialectal variation. Of the four varieties considered in this chapter, Riau Indonesian has inclusory pronouns, Jakarta Indonesian has associative plurals, Kupang Malay has both inclusory pronouns and associative plurals, while Papuan Malay has a single construction with properties of both an inclusory pronoun and an associative plural. We now consider each of the four dialects in turn.

In Riau Indonesian there is an inclusory pronoun construction, as in *kami sama Adi* (1PL.EXCL together Adi) 'me/us and Adi', where *sama* is a grammatical marker with a wide range of functions including, among others, 'with', 'and', 'same', and a marker of various non-absolutive thematic roles; see Gil (2004) for detailed description and analysis of *sama*. Occasionally, instead of *sama*, the more semantically specific *dengan* 'with' may appear. While *kami* is the only pronoun that may enter into the inclusory construction, the second term may contain either a proper or a common noun, in which latter case it may be either lexical or phrasal. In a more semantically specific variant of the construction, the marker *sama* (or *dengan*) may be replaced by the numeral *dua* 'two', resulting in what might be referred to as a dual inclusory pronoun construction, for example *kami dua Adi* (1PL.EXCL two Adi)

'me and Adi'. However, higher numerals may not occur in this position.[11] It should be noted, though, that alongside the inclusory pronoun construction, Riau Indonesian also allows the option of a regular, non-inclusory coordination of a pronoun with a proper noun, as in *aku sama Adi* (1SG together Adi) 'me and Adi'.

In contrast, in Jakarta Indonesian there is no inclusory pronoun construction; however, there is an associative plural construction, as in *Adi cs* (Adi friend) 'Adi and associate(s)'. The form *cs*, pronounced [cees], derives ultimately from the Latin associative plural marker *cum suis*; however, unlike in Latin, *cs* may occasionally stand on its own, where it simply means 'friend'. The associative plural construction with *cs* is clearly grammaticalized, as it occurs with no intervening conjoining element. In this respect, it contrasts with more productive and compositional phrases involving the usual word for friend, *temen*, such as *Adi (a)ma temennya* (Adi together friend:ASSOC) 'Adi and his friend(s)', in which the conjoining element *(a)ma* is obligatory, and *temen* 'friend' is suffixed with the associative *-nya*, which marks its host as being associated with some other entity, which, in the present context is understood as the possessor *Adi*.

Moving east across the archipelago the plot thickens. In Kupang Malay there is an inclusory pronoun construction similar to the Riau Indonesian one, for example *botong deng Yuda* (1PL.EXCL with Yuda) 'me/us and Yuda'. However, in addition, there is also an associative plural construction, formed with the 3rd person plural *dong*, as in *Yuda dong* (Yuda 3PL) 'Yuda and associate(s)'. As shown in Daniel (2000), 3rd person plural pronouns are a cross-linguistically widespread means for the formation of associative plurals; some other languages employing this strategy include Mandarin, Hausa, and even some dialects of English, such as Texas English *John 'em*.

Associative plurals in Kupang Malay bear a close affinity to two other related constructions. In the first construction, the human proper-noun focal referent of the associative plural is replaced by a non-human noun. In many cases, the noun in question is a proper noun, for example *Oebobo dong* 'the Oebobo guys' (Oebobo is the name of a neighborhood), *Awu dong* 'the Awu guys' (Awu is the name of a ship). In other cases, though, it is a common noun, or even a phrasal expression, for example *be pung kampung dong* (1SG POSS village 3PL) 'the people from my village'. Such constructions differ from associative plurals in that their reference does not include the focal element. We shall refer to these constructions as *extra-focal associative plurals*.[12] Similar constructions involving toponyms plus a plu-

---

11 In contrast, in Coastal Lampung Indonesian, the numeral *tiga* 'three' may also occur in an inclusory construction, for example *kitorang tiga Adi* (1PL.EXCL three Adi) 'us two and Adi'. I do not know whether higher numerals are also possible.

12 It would appear that there is no commonly accepted term for these constructions in the linguistic literature. In principle it would seem advisable to broaden the reference of *associative plural* to encompass two subtypes, focal associative plurals, representing the familiar construction, and then the extra-focal associative plurals introduced above. However, readers reluctant to make the termi-

ral marker have been described for the Uralic language Mari by Daniel (1999:369, 2000, 2019).[13]

In the second construction, the pronoun *dong* is denuded of its associative semantics and reinterpreted as a simple marker of additive plurality; in this construction, *dong* is simply a plural word. While the use of *dong* is optional, it occurs quite frequently and in a wide range of environments, as illustrated in the following naturalistic utterances:

(6)   *Paling   satu dua kata   dong bagitu*
       SUPERL one two word 3PL   like:DEM:DEM.DIST
       'At most one or two words.'
       [Speaker explaining how she can understand when her parents speak Ende but not speak it herself.]

(7)   *Bap   Besa pung jadwal   dong kaco   balo   abis*
       father uncle POSS schedule 3PL   disturb MUT\disturb finish
       'My schedule is all topsy-turvy.'
       [Bap Besa complaining about how busy he is.]

In both (6) and (7), *dong* pluralizes an inanimate word, *kata* 'word' and *jadwal* 'schedule' respectively. In (6), *dong* applies to a word that is already numerically quantified, and hence its pluralizing function is redundant. In contrast, in (7), *dong* "looks inside" the collective noun *jadwal* 'schedule', underscoring the plurality of its constituent parts, or 'schedule items', in a way reminiscent of the individualization of collective nouns such as *tentara* 'army' > 'soldier' and *polisi* 'police' > 'policeman/policewoman' discussed in Section 2.3.2 above.

The contrast between associative and additive plural functions of Kupang Malay *dong* is illustrated in the following minimal pair consisting of two naturalistic utterances containing the expression *oto dong* (car 3PL):

(8)   *Oto dong hanya batareak*
       car 3PL only NON.PAT:shout
       'The car drivers just shout.'
       [Speaker describing her trip.]

---

nological leap are invited to interpret the adjective *extra-focal* as non-restrictive, that is to say, to consider extra-focal associative plurals as constituting a construction that does not fall within the domain of "true" associative plurals.

13 A subtly different construction in West Greenlandic is described by Fortescue (1984:247) and subsequently referred to as *anti-associative* by Corbett (2000:241). In West Greenlandic, the plural form of the word for 'ship', *umiarsuit*, receives the interpretation 'ship plus its crew'; this construction thus differs from the Kupang Malay extra-focal associative plural in that the referent of the host word, namely the ship, is included in the reference of the entire expression, whereas in Kupang Malay it is excluded.

(9) *Oto dong su   ponuh samua*
    car 3PL   PFCT full   all
    'The cars are already all full.'
    [Speaker describing her trip.]

In (8), *oto dong* is an extra-focal associative plural, referring to the people associated with the cars. In contrast, in (9), *oto dong* is a regular plural, referring simply to the cars. The contrast between the two different constructions is effected by the different selectional restrictions associated with their respective predicates: only people can shout, while only cars can be full. As pointed out by Maurer et al (2013), the co-opting of a 3rd person plural pronoun for the marking of ordinary additive plurality is quite common cross-linguistically, especially in creole languages, for example Cameroon Pidgin English.

As for the identity of associative and additive plurals displayed by Kupang Malay, this is also cross-linguistically widespread, found in 104 out of 200 languages with associative plurals in the Daniel and Moravcsik (2005) 237-language sample, for example Yoruba, Turkish and Nias. However, while the associative plural would seem in many cases to have developed from an ordinary additive plural, the Kupang Malay case appears to have possibly followed a rather different trajectory, from pronoun to associative plural marker / extra-focal associative plural marker to additive plural marker. Some insight into how such grammaticalization might have occurred derives from utterances which may be analyzed as providing bridging contexts between the two functions, such as the following:

(10) *Oto dong antar   sampe SMA   lima sana*
     car 3PL   accompany arrive senior.high.school five there
     'The car drivers take us to SMA 5.' / 'The cars take us to SMA 5.'
     [Speaker discussing transportation options.]

In the above utterance, the context is such that the expression *oto dong* could equally readily be understood either as an extra-focal associative plural referring to the car drivers, as in (8), or as an additive plural referring to the cars themselves, as in (9). Example (10) thus suggests a potential pathway via which the additive plural function might have arisen out of an earlier extra-focal associative plural.

Moving on now to Papuan Malay, we first note that there are no inclusory pronoun constructions of the kind present in Riau Indonesian and Kupang Malay. As for constructions with the 3rd person plural pronoun *dong,* we observe that strings such as *Lorens dong* 'Lorens and associate(s)', *Dorondola dong* 'the Dorondola guys' (Dorondola is also the name of a ship) work just like their counterparts in Kupang Malay. From a comparative typological perspective, then, one might conclude that Papuan Malay has associative plurals and extra-focal associative plurals similar to those in Kupang Malay. However, such superficial etic similarities mask deeper underlying emic differences.

To begin with, in Papuan Malay, the use of *dong* as an ordinary additive plural word is unavailable for inanimate nouns. Thus, whereas in Kupang Malay, the string *oto dong* (car 3PL) can mean either 'people associated with cars' as in (8), or just 'cars' as in (9), in Papuan Malay, the corresponding string, *mobil dong*, can only mean 'people associated with cars'. As an additive plural marker, *dong* occurs only with animate nouns, as in *guru dong* (teacher 3PL) 'teachers', *anjing dong* (dog 3PL) 'dogs'. This would seem to suggest that in such constructions, the pronoun *dong* maintains its pronominal nature, and in particular, its specification as animate (in contrast to its inanimate counterpart *akan*, cf. Table 4 above).

In fact, the pronominal nature of *dong* in the above constructions is part and parcel of a more far-reaching difference between Kupang and Papuan Malay dialects pertaining to the pronoun itself. While in Kupang Malay, *dong* can only be replaced by other variants of the 3rd person plural, in Papuan Malay, *dong* can be replaced by most of the pronouns listed in Table 4. Table 5 below illustrates the resulting paradigm in the context of the proper noun *Lorens*:

Tab. 5: Proper Noun plus Pronoun in Papuan Malay.

	singular	plural	dual
1st person	*?Lorens saya [sa]*	*Lorens torang [tong]*	*Lorens tendua*
2nd person	*Lorens koi [ko]*	*Lorens kamu [kam]*	*Lorens kamdua*
3rd person	*Lorens dia [de]*	*Lorens dorang [dong]*	*Lorens dendua*
3rd person inanimate	**Lorens akan*		

As shown in Table 5, all of the combinations of noun plus pronoun are grammatical, with the exception of the 3rd person inanimate *akan*, and possibly also the 1st person singular *saya [sa]*. But what do these constructions mean? The answer is actually straightforward: in all of the forms in Table 5, the pronoun specifies a set of referents with particular number and person values, while the proper noun asserts that the person to whom it refers is included in the specified set. For example, in an expression such as *Lorens tendua* (cf. the similar *Empong tendua* in example (4) above), the pronoun specifies a set of referents with 1st person dual reference, while the proper noun asserts that the specified set must include Lorens, resulting in the interpretation 'me and Lorens'.

For the singular, the pronoun is completely coreferential with the preceding proper noun; the pronoun is thus truth-conditionally redundant, and would seem to serve some kind of discourse function; Donohue and Sawaki (2007) refer to this function as that of a specifier. Some examples of this construction from written social media are given in (11) and (12) below:

(11) *Tra usa krm uda ...? Z suh malas den pit ko*
 Tra usa   kirim suda Sa su   malas   deng Pit ko
 NEG must send PFCT 1SG PFCT reluctant with Pit 2SG
 'Don't bother sending it any more. I'm already tired of you.'
 [SMS message: speaker angry that I hadn't sent him any phone credit.]

(12) *oky piiitt ... Nanti ten2 sama sama jln cari papua dia*
 Oke Pit nanti   tendua sama-sama   jalan cari   Papua dia
 okay Pit FUT.PROX 1DU   DISTR~together go   search Papua 3SG
 'Okay David, later we'll go together and look for Papua.'
 [Facebook chat: speaker offering to help me look for a friend called Papua
 who had gone missing.]

In (11) and (12), the phrases *Pit ko* and *Papua dia* illustrate the contextually redundant marking of proper nouns for singular number, as well as for person, 2nd and 3rd respectively.

In contrast, for the plural and dual forms, the pronouns are no longer redundant, but rather add further information pertaining to the number and person of the expression. From an etic perspective, the 3rd person forms *Lorens dong* 'Lorens and associate(s)' and *Lorens dendua* 'Lorens and associate' would seem, as noted above, to constitute garden-variety instantiations of the comparative concepts of associative plural and associative dual respectively. But what, then, of the 1st and 2nd person forms? Most commonly, associative plurals (and duals) are 3rd person; however, the 1st and 2nd person forms in Table 5 appear to suggest that the notion of associative plural (and dual) should also include expressions of 1st or 2nd person reference. Conversely, maintaining the etic perspective, the 1st person forms *Lorens tong* 'me, Lorens and associate(s)' and *Lorens tendua* 'me and Lorens' would appear to be typical examples of the comparative concept of inclusory pronoun. However, while inclusory pronoun constructions are often limited to 1st person, as in Riau Indonesian and Kupang Malay above, or at least tend to occur more frequently in 1st person, the 2nd and 3rd person forms in Table 5 would appear to suggest that the notion of inclusory pronoun should be understood as admitting also expressions of 2nd and 3rd person reference.[14]

---

14 Indeed, with regard to 3rd person, a similar argument could be made for many of the numerous languages reported to have an associative plural construction formed with a 3rd person plural pronoun. Specifically, the argument would go through for languages in which, as is the case in Papuan Malay, the associative plural construction allows for the possibility of exactly two referents, thereby entailing that one of the two referents must be that of the host proper noun. In contrast, in other languages with an apparently similar construction, such as the Texas English *John 'em*, the two-referent interpretation is unavailable, thereby suggesting that they are not also instances of an inclusory pronoun construction.

So what are the plural and dual forms in Table 5: inclusory pronouns or associative plurals/duals? From an etic typological perspective, some are obviously inclusories, while others are equally clearly associative plurals/duals; indeed, depending on how liberally we wish to interpret these two comparative concepts, many of the forms may be said to instantiate both construction types. On the other hand, some of the other forms, specifically the singular ones exemplified in (11) and (12) above, are clearly neither inclusory pronouns nor associative plurals/duals. Thus, the etic typological perspective has the effect of carving Table 5 up into a handful of very different albeit possibly overlapping regions instantiating distinct construction types. However, from an emic language-internal perspective, there is no evidence whatsoever for such a partitioning of the forms in Table 5. Rather, the entirety of Table 5 clearly instantiates a single language-specific construction with a unified semantic description, in which the pronoun specifies a set of referents, one of which is also referred to by the host expression. We shall accordingly refer to the forms in Table 5 simply as P(erson)-N(umber)-marked expressions.

PN-marked expressions in Papuan Malay are thus a construction consisting of a host expression followed by a pronoun. Whereas in Table 5, the host is a proper noun, in other cases, the host may belong to other categories, with concomitant adjustments to the resulting semantics. Examples (13) and (14) (the latter containing a bit of code-switching with English) illustrate PN-marked expressions in which the host is a common noun, the address term *pace* 'male friend':

(13) *sapa yg pncuri pace ko punya pls k ....??? rumput-rumput dong kapa*
    Siapa yang pancuri pace   ko punya pulsa   ka
    who REL AG:steal male.friend 2SG POSS phone.credit Q
    Rumput-rumput dong kapa
    DISTR~grass   3PL   Q:what
    'Who stole your phone credit? Those bums or what?'
    [Facebook chat: speaker responding to my complaint about missing phone credit.]

(14) *boas yang andalan ka ..... the best striker di indonesia .... siapa yang mo lawan pace dia ....*
    Boas yang andalan ka The best striker di Indonesia Siapa yang mo
    Boas REL cool   Q   LOC Indonesia who REL want
    lawan pace   dia
    oppose male.friend 3SG
    'Cool Boas huh, the best striker in Indonesia, who would want to go up against him?'
    [Facebook status update: discussing football.]

Examples (13) and (14) illustrate the PN-marked expressions *pace ko* and *pace dia*. Unlike proper nouns, common nouns such as *pace* are unspecified for number.

Hence, while in (11) and (12), with proper nouns *Pit* and *Papua,* the number-marking function of the pronoun is truth-conditionally redundant, in (13) and (14) above the singular pronouns impose singular interpretations on the PN-marked expressions – from an etic perspective, they might be considered as a species of singulative marker. Conversely, when the common noun is in construction with a plural or dual pronoun, the pronoun assumes its role as a kind of plural or dual marker. Thus, whereas *Lorens dong and Lorens dendua* both contain a focal referent, and hence qualify as associative plurals/duals, *pace dong* and *pace dendua* do not contain a focal referent, and therefore are more appropriately considered as simple additive plurals. From an emic language-internal perspective, then, the additive plural-marking function of Papuan Malay pronouns thus emerges as a particular case of their use in the more general PN-marking construction.

Further semantic adjustments are in evidence in the case of inanimate hosts, for which the animate semantics of the pronoun precludes the host from being included in the set of referents specified by the pronoun. For example, for proper nouns such as locative *Borarsi* (the name of a neighborhood), the resulting expressions assume extra-focal associative interpretations, such as *Borarsi dia* 'the Borarsi guy', *Borarsi kam* 'you Borarsi guys'. In the case of common nouns such as *mobil* 'car,' similar extra-focal associative interpretations are also available, *mobil dia* 'the car guy', *mobil kam* 'you car guys', as well as the previously mentioned *mobil dong* 'the car guys'.

Finally, in some cases, the host of the pronoun may be some kind of adjectival or verbal phrase; in cases such as these, the pronoun marks the number and person of a set of referents associated, in the role of theme or participant, with the host expression. For example, in (13) above, the expression *rumput-rumput dong* (DISTR~grass 3PL) 'those bums' refers to a set of entities constituting the theme of *rumput*, a slang expression for down-and-out bums. Similarly, in expressions such as *minum balo dia* (drink toddy 3SG) 'the toddy-drinking guy', *minum balo kam* (drink toddy 2PL) 'you toddy-drinking guys', and so forth, reference is to a (singular or plural) set of entities constituting the agent of *minum balo*.

Thus, PN-marked expressions in Papuan Malay constitute a productive means for marking number, as well as person, in a wide variety of morphosyntactic environments. In all of these environments, the pronoun specifies a set of referents with particular number and person values, where one or more of the referents may be coreferential with the host expression. From an etic cross-linguistic perspective, PN-marked expressions encompass a range of comparative concepts, including, among others, singulatives, ordinary additive plurals, associative plurals, extra-focal associative plurals, inclusory pronouns and possibly others. However, the profusion of comparative concepts available to the typologist obscures the underlying unity of the PN-marking construction as emerging from an emic, language-internal analysis.

Conversely, the present discussion shows how superficially similar constructions in closely related dialects may actually instantiate very different structures.

For example, in both Kupang and Papuan Malay, a proper noun followed by the 3rd person plural pronoun *dong* gets translated into English as 'so and so and his/her associate(s)'. However, whereas in Kupang Malay *Yuda dong* represents a cross-linguistically rather typical case of an associative plural formed with a 3rd person plural pronoun, in Papuan Malay *Lorens dong* is an instance of a much more general construction, that of PN-marked expressions, endowed with a substantially wider range of morphosyntactic manifestations and corresponding interpretations.

2.4 Verbal number

Like their nominal counterparts, verbs in Indonesian are unspecified for number. Unlike nouns, however, verbs have the potential of expressing number in two conceptually distinct dimensions, pertaining, respectively, to the quantity and number of activities denoted by the verb, and to the quantity and number of participants associated with each of the verb's arguments. The former dimension falls largely within the domain of aktionsart and aspect; while aktionsart is largely inherent to the lexical semantics of the verb, aspect is marked by a small number of optional periphrastic markers which lie beyond the scope of the present survey. The latter dimension, though, may be optionally expressed by a number of morphosyntactic devices whose function is to specify the number of one of the verb's arguments.

2.4.1 The plural word *pada*

In Jakarta Indonesian, though in none of the other dialects considered here, a form *pada* may occur in front of its host, marking it as having a plural argument. Following are four examples of *pada*, occurring in front of hosts *ngapain* 'do what', *gede* 'big', *sekola* 'school', and *di luar* 'outside':

(15) *Lagi pada ngapain   sih?*
PROG PL   AG:what:EP EMPH
'What are they doing?'
[Speaker playing with child, asking about a couple of dolls.]

(16) *Karang mah uda pada gede*
now   TOP PFCT PL   big
'They're all grown up now.'
[Speaker complaining about some disobedient young adults.]

(17) *Kalo misalnya   uda jam delapan, orang pada sekola*
TOP example:ASSOC PFCT hour eight   person PL   school
'When it was say eight o'clock, people would be going to school.'
[Speaker telling a story about her neighborhood.]

(18) *Ya   mungkin yang punya-punya warung itu*   *pada di   luar*
     yes maybe   REL DISTR~own kiosk DEM:DEM.DIST PL   LOC outside
     'Yes, maybe the ones who owned the kiosks were outside.'
     [Speaker telling a story about her neighborhood.]

As suggested by the above examples, *pada* may occur with a wide variety of hosts, in keeping with the monocategorial nature of Indonesian grammar. Nevertheless, the hosts of *pada* are invariably interpreted as denoting activities or properties. Thus, in (17) *sekola* is interpreted as 'go to school', while in (18) *di luar* is understood as 'be outside'. The argument marked by *pada* as plural need not be overtly present; while in (17) and (18) it is, namely *orang* and *yang punya warung itu* respectively, in (15) and (16) it is not. Finally, the marker *pada* is itself always optional; the same utterances would permit the same plural-argument interpretations if *pada* were absent, though of course they would then also allow additional interpretations in which the argument in question were singular.

   Examples (15)–(18) would seem to suggest that the argument marked by *pada* as being plural is that which might be considered to be its host's subject. Such a conclusion is supported by speakers' judgements of the following artificially constructed example:

(19) *Anjing lagi   pada kejar kucing*
     dog   PROG PL   chase cat
     (a)   'The dogs are chasing the cat.'
     (b) *'The dog is chasing the cats.'

As suggested by (19), *pada*, here hosted by the transitive verb *kejar*, may mark its subject *anjing* as plural, as in (19a), but not its object *kucing*, as in (19b). (Of course, since nouns are unmarked for number, *kucing* could still happen to be plural, but this would not be because of the presence of *pada*.) However, the apparent subject-orientation of *pada* is belied by the following examples involving *pada* hosted by a verb marked with the patient-orientation prefix *di-*:[15]

(20) *Ya, asyik, sudah pada diwarnain*
     yes great PFCT PL   PAT:colour:EP
     'Yeah, great, you've coloured lots of them.'
     [Speaker talking to child who has already finished colouring several parts
     of an elephant picture.]

---

15 While the prefix *di-* is usually considered to be a passive marker, Gil (2002) and Conners, Bowden and Gil (2015) show that in Jakarta Indonesian as well as other colloquial varieties of Indonesian, it lacks certain criterial properties of a passive marker, suggesting that is more appropriately analyzed as a "generalized patient-oriented voice marker".

(21) *Sekarang kan pada dibilangnye   ape?*
　　 now   Q   PL   PAT:say:ASSOC what
　　 'Now what do we call it?'
　　 [Speaker talking about a type of traditional Betawi theatre that used to be
　　 called *lenong.*]

(22) *Pada dibikin   rumah, ya?*
　　 PL   PAT:make house   yes
　　 'Use them to make a house, okay?'
　　 [Speaker and infant playing with Lego, speaker draws all the Lego pieces
　　 together and presents them to child.]

In (20), the argument marked by *pada* as plural is the parts of the elephant picture, or the patient of *diwarnain*, which, under a passive analysis, would be its surface subject. In contrast, in (21), the argument marked by *pada* as plural is the people who call the theatre by a certain name, or agent of *dibilangnye*, which, under a passive analysis, would be its underlying but not surface subject. Finally, in (22), the argument marked by *pada* as plural is neither the agent of *dibikin*, the child, nor its patient, the house, but rather its instrument, the Lego pieces. Thus, as suggested by (20)–(22), the choice of argument pluralized by *pada* is quite variable; however, as shown in (19) previously it is not completely unconstrained. At present, I have no account for what the appropriate generalizations governing the selection of pluralized argument might be.

In constructions containing *pada*, the pluralized argument, when present, may assume a variety of forms: 1st, 2nd or 3rd person pronouns, a conjunction of proper nouns, and so forth. Of interest to us here are the number properties of the argument expression when considered on its own. In (17), *orang* is unspecified for number, so the effect of *pada* is to restrict its interpretation to plural. In contrast, in (18), *yang punya-punya warung itu* is already marked as plural by reduplication (see Section 2.4.2 below), and hence the plural-marking function of *pada* is truth-conditionally redundant. However, in yet other cases, the argument pluralized by *pada* may actually "start off as" singular.

In (23) and (24) below, *pada* is oriented towards the singular expressions *ibu saya* 'my mother' and *Dada*, a proper noun here being used with 1st person reference:

(23) *Kan dulu ibu   saya pada di   situ,   sudara di*
　　 Q   first mother 1SG PL   LOC LOC:DEM:DEM.DIST sibling LOC
　　 situ
　　 LOC:DEM:DEM.DIST
　　 'A long time ago my mother and her relatives were there.'
　　 [Speaker reminiscing about where they lived.]

(24) *Nyang sekarang ruma makanan yang Dada suka pada pegi tu*
REL now house eat:AUG REL Dada like PL go DEM.DIST
'It's now a restaurant that we like to go to.'
[Dada talking about her neighborhood.]

In cases such as these, *pada* imposes an associative plural interpretation on the relevant argument. In (23), *ibu saya* constitutes the focal element for the interpretation 'my mother and associates', which in this case is further explicated in the continuation of the utterance by the expression *sudara* 'sibling'. And in (24), *Dada* forms the focal element for the interpretation 'me and associates', or, more simply, 'we'. Since *pada* and its selected argument do not form a constituent, constructions such as these may be referred to as *discontinuous associative plurals*. A further and somewhat different example of a discontinuous associative plural construction in Papuan Malay is provided in the Section 3 below, see example (58b).[16]

A rather different kind of interpretation is obtained in (25) and (26), also involving *pada* being oriented towards a singular argument, *Jalan Kalisari* 'Kalisari street' and *perut aku* 'my stomach' respectively:

(25) *Trus, Jalan Kalisari itu masi belom pada diaspal*
continue street Kalisari DEM:DEM.DIST still NEG.PFCT PL PAT:asphalt
si dulu situ
EMPH first LOC:DEM:DEM.DIST
'So Kalisari street was still unpaved back then there.'
[Speaker reminiscing about where they lived.]

(26) *perutt ku dh pda demo nee*
Perut aku dah pada demo ni
stomach 1SG PFCT PL demonstration DEM.PROX
'My stomach's all rioting.'
[Speaker chatting on facebook, complaining about how hungry he is.]

In the above examples, *pada* imposes a *mereological* interpretation on the selected argument, drawing attention to its partitioning into a plurality of parts. In (25), *pada* emphasizes the extent of the street and the completeness of its ultimate paving.[17] And in (26), *pada* reinforces a vivid metaphorical description of the speaker's stom-

---

16 Variants of the discontinuous associative plural construction have been described by several authors, for example Harbour and McKenzie (this volume: 724), Fabri (1993:276–278) for Maltese, Corbett (2000:191–192) for the Talitsk dialect of Russian, and Dalrymple and Mofu (2012:31) for the Austronesian language Dusner.
17 Under an alternative interpretation, *pada* might also be referring to the plurality of the workmen who paved the road, however the mereological interpretation would appear to be the dominant one.

ach, depicted as playing host to a riot; here, the semantic effect of *pada* might perhaps be conveyed in English as 'all over it'.

2.4.2 Reduplication

Reduplication, in its plurality-marking function, is perhaps one of the most renowned grammatical features of Indonesian; see, for example, Chung (2000) and Lander (2003). Nevertheless, many of the salient characteristics of reduplication in Indonesian remain underdescribed. While reduplication is common in all four varieties of Indonesian under consideration here, impressionistically its range of forms, variety of functions, and textual frequency all tend to decrease from west to east. This section focusses on reduplication in the westernmost of the four varieties, and hence that in which it is arguably at its most productive: Riau Indonesian.

Most commonly, reduplication applies to a whole word, creating an additional copy thereof; as shown in Gil (2020), the input to reduplication is typically a disyllabic morpheme plus one or more optional bound morphemes, though in a few cases, longer strings containing two or more disyllabic morphemes may also be reduplicated. In a few instances, reduplication is partial, involving just the first CV-sequence of the word, e.g. in (39) and (50) below. Also in a few cases, reduplication is multiple, creating a sequence of three or more occurrences of the reduplicated material, e.g. in (47) below. The choice between these alternative formal strategies would seem to bear no grammatical or semantic consequences; at most, partial and multiple reduplication may sometimes be associated with greater expressivity. Reduplication is highly productive, and may apply to pretty much any content word, as well as to most or all function words, including pronouns, demonstratives and interrogatives; however, its use is always optional, in no cases is it required by a particular grammatical environment. As argued in Gil (2005a), the absence of a clear cut distinction between morphology and syntax sometimes makes it difficult, in some cases, to distinguish between reduplication and repetition; however, nothing in the following discussion depends crucially on being able to keep the two apart.

On the face of it, reduplication is associated with a very wide range of functions. Nevertheless, many of these functions appear straightforwardly derivable from the notion of plurality; among these are plurality of kinds, plurality of locations, plurality of arguments, and plurality of actions, or iterativity. A further set of closely related functions involves quantification over entities lacking in individuation; these include large size of an object, large extent of time, that is to say durativity, and intensification of a scalar property. However, many other functions of reduplication do not appear to be related to the notion of plurality; among these are atelicity, negative polarity, concessivity and deprecation. Complicating the analysis of reduplication is the fact that in many utterances in which it occurs, it would appear to be fulfilling more than one of the above functions simultaneously. Accordingly, in

order to justify the existence of the above functions, it is first necessary to examine utterances in which a single function can be isolated, as in (27)–(41). Only after that is it possible to consider utterances in which two or more of the functions are coexpressed, as in (42)–(54).

The association of reduplication with plurality is extremely common cross-linguistically; see Moravcsik (1978) and others. However, in Riau Indonesian, utterances in which reduplication expresses nothing but simple plurality of objects are considerably less frequent than one might perhaps have expected. Following is one such example:

(27) *Anak-anak   singa*
     DISTR~child lion
     'Lion cubs.'
     [Watching nature program on TV, speaker commenting on picture of lion cubs feeding from mother.]

Occasionally, reduplication may instead express large size:

(28) *Ini   kan pulau-pulau   Padang*
     DEM:DEM.PROX Q   DISTR~island Padang
     'That's all Padang island.'
     [On pier facing strait, speaker points to the island of Padang island, stretching for almost 180 degrees from left to right, explaining to visitor.]

In the above example, the large-size interpretation, expressed by reduplication, may be derived from plurality by mereological quantification, in that the island is visible in a large number of different directions.

Given that Riau Indonesian does not distinguish grammatically between nouns and verbs, it is not surprising to find precise counterparts to (27) and (28) in which quantification is over activities rather than objects:

(29) *Tadi   nanya-nanya udah sampai belum*
     PST.PROX DISTR~AG:ask PFCT arrive   NEG.PFCT
     'Earlier he kept on asking whether you had already arrived.'
     [Speaker to guest who has arrived in village to visit friend, informing him that his friend is out.]

(30) *Lama   kali, saya capek nunggu-nunggu*
     long.time very 1SG tired   DISTR~AG:wait
     'Such a long time, I was tired of waiting.'
     [Speaker berating friend for being back late.]

Example (29) illustrates the very common function of reduplication as a marker of iterativity; as in (27), here too reduplication marks plurality, the only difference is that it is counting activities rather than objects. And example (30) shows the equally common function of reduplication as marking durativity; again, as in (28), reduplication is expressing largeness of size, understood as the extent of an activity. Although nominal and verbal quantification are often considered to be distinct phenomena, the parallel between the above two pairs of examples suggests that from a language-internal perspective, there is no reason, grammatical or semantic, not to subsume the above four functions, plurality, large size, iterativity and durativity, under a more generalized notion of plurality.

Indeed, such a more generalized notion of plurality may be further extended to encompass several other distinct functions expressed by reduplication. One important function, exemplified in (31)–(34) below, is plurality of arguments:

(31) *Ayo,   penumpang turun-turun   semuanya*
EXHRT passenger   DISTR~descend all:ASSOC
'All passengers off.'
[Minibus conductor to passengers, having arrived at destination.]

(32) *Damsir beli celana   sama   si   Man   sudah bulu-bulu*
Damsir buy trousers together PERS FAM\Mansudir PFCT   DISTR~feather
'The pair of trousers that Damsir bought with Mansudir are already frayed all over.'
[Speaker talking about his friends' clothes.]

(33) *Ngapain   kalian nengok-nengok orang   minum?*
AG:what:EP 2PL   DISTR~AG:look person drink
'Why are you watching the person drink?'
[Food stall owner to children gathered around me as I drink.]

(34) *Saya masuk-masukkan semua*
1SG   DISTR~enter:EP   all
'I'll pocket them all.'
[Playing billiards.]

In (31), reduplication of *turun* implies the plurality of its theme argument, *penumpang*. Similarly, in (32), reduplication of *bulu* entails the plurality of its theme argument, *celana*, except that here, quantification is mereological, as in (28) earlier. Examples (33) and (34) illustrate reduplication of a transitive verb: whereas in (33), reduplication of *nengok* marks its agent *kalian* as plural, in (34), reduplication of *masukkan* marks its understood patient, the billiard balls, as plural. In conjunction, then, (33) and (34) show that the choice of argument pluralized by reduplication is

unconstrained. In this respect, Riau Indonesian reduplication differs from Jakarta Indonesian *pada*, which, as shown in example (19) earlier, cannot mark the patient of a transitive verb as plural. More generally, it differs also from verbal number marking in many other languages, where it is typically constrained in accordance with a variety of grammatical and/or semantic hierarchies.

Prima facie, examples (31)–(34), in which reduplication marks plurality of arguments, seem rather different from those in (27)–(30), in which reduplication expresses various guises of a more basic plurality. However, by generalizing the notions of semantic frame and argument structure to include the role of *essant*, introduced in Gil (2013b), it is possible to subsume the examples in (27)–(30) under a broader notion of plurality of arguments. The essant role is that which is cross-linguistically prototypically instantiated by the subject of a predicate nominal construction. For example, in (27) above, *anak-anak singa* assigns the role of essant to the objects being viewed on the TV screen, that is to say to the 'They' in a more elaborate potential English translation 'They are lion cubs'. Thus, just as *masuk-masukkan* in (34) marks its understood patient, the billiard balls, as plural, so *anak-anak singa* in (27) marks its understood essant as plural, where in the case at hand the essant refers simply to the objects that *are* lion cubs. Analogous analyses are available also for (28)–(30); for example *nanya-nanya* in (29) also marks its understood essant as plural, in the present case the activities that *are* askings. Thus, in all of the examples of reduplication considered so far, involving ordinary plurality, large size, iterativity, durativity, and the plurality of roles such as theme, agent and patient, reduplication may be analyzed in unified fashion as marking the plurality of a contextually selected argument.

In accordance with the above analysis, the reduplicated expression may actually be more insightfully understood as a marker of distributivity. In accordance with Gil (1982, 1988, 1992), Choe (1987) and others, distributivity is a binary relationship holding between two terms, the *distributive key* and the *distributive share*. The distributive key is semantically plural, each of its members being associated with its respective distributive share. For example, in an English sentence such as *Each boy got three pencils*, the subject *each boy* is distributive key while the object *three pencils* is distributive share: each member of a plural set of boys is associated with three pencils. In Riau Indonesian, in all of the above examples, reduplication marks its host as distributive share, thereby implying the presence of some other term as its plural distributive key. In other words, the reduplicated expression "distributes over" another term which receives a plural interpretation. Revisiting some of the preceding examples, in (27) *anak-anak*, the property of being a lion cub distributes over a set of objects that are lion cubs, in (29) *nanya-nanya*, the property of being an asking distributes over a set of sequential activities that are askings, while in (34) *masuk-masukkan*, the property of being pocketed distributes over a set of billiard balls.

Cross-linguistically, as shown in Gil (1982, 1988, 1992), numerals constitute the most common host for the marking of distributivity, and indeed, in Riau Indonesian,

reduplicated numerals are commonly understood as distributive-share, as in the following example:

(35) *Isinya   empat-empat orang?*
     contain:ASSOC DISTR~four   person
     'Do they take four people each?'
     [At tourist resort, speaker asking about fishing boats.]

In (35) above, *empat-empat*, in conjunction with its head noun *orang*, constitutes the distributive share, and the understood boats its distributive key: each boat takes four people. Thus, distributive numerals such as *empat-empat* provide support for the unified analysis of reduplication accounting for all of the cases in (27)–(35), in accordance with which reduplication marks its host as distributive share, and seeks out another semantically plural entity, present or implied, as its distributive key.

While accounting for a wide range of superficially quite different constructions, a number of additional constructions involving reduplication are not straightforwardly amenable to an analysis involving distributivity. Indeed, one such construction, illustrated below, also involves numerals:

(36) *Dengarkan dua-dua*
     listen:EP   DISTR~two
     'Listen with both of them.'
     [Speaker handing walkman to interlocutor, telling him to use both earphones.]

Unlike (35), here *dua-dua* is not obviously distributive share; the understood earphones number two in total, not two for each member of some other semantically plural set of entities functioning as distributive key. On the face of it, reduplication here would seem to be marking its host as a so-called "collective numeral".[18]

Another somewhat different notion expressed by reduplication is that of intensification of scalar properties:

(37) *Bikin panjang-panjang*
     make DISTR~long
     'Make it very long.'
     [Playing a computer billiards game in which the length of the cue is controllable, determining the strength of the shot, onlooker advises one of the players.]

────────

18 A potential strategy for subsuming collective numerals such as *dua-dua* within the above-proposed analysis of distributivity might involve characterizing the individual items, here the single earphones, as the plural distributive key, and the property of being part of a collective set, here numbering two items, as the distributive share, marked by reduplication. That is to say, each individual earphone, the distibutive key, belongs to a set of cardinality two, the distributive share. Such

Intuitively, intensification seems to be semantically akin to large size and durativity, though it is not immediately obvious how examples such as (37) can be analyzed in terms of a relation of distributivity.[19]

A number of additional functions associated with reduplication would appear to be unrelated to distributivity. One very common function of reduplication is atelicity:

(38) *Minum-minum dulu*
　　 DISTR~drink 　first
　　 'Have a drink first.'
　　 [Minibus conductor suggesting to passenger that he have drink while waiting for vehicle to depart.]

Another common function is negative polarity, in which a reduplicated expression occurs within the scope of a marker of negation:

(39) *Aku tak ada du-duit, 　　 bang*
　　 1SG　NEG　exist　DISTR~money　FAM\elder.brother
　　 'I don't have any money.'
　　 [Speaker trying to borrow money from friend.]

Yet another function of reduplication is concessivity:

(40) *Kecil-kecil 　 berat e, kecil-kecil 　 tapi berat*
　　 DISTR~small　heavy　Q　DISTR~small　but　heavy
　　 'Small but heavy, small but heavy.'
　　 [Speaker talking about a laptop computer.]

And finally, reduplication can be used to express an attitude of deprecation:

(41) *Entah 　 lah 　　 topi-topi 　 yang itu*
　　 UNCERT　NEG.FOC　DISTR~hat　REL　 DEM:DEM.DIST
　　 'I don't know anything about that hat.'
　　 [On long distance bus, about to depart after rest stop; a passenger calls out that somebody's left their hat behind; and conductor responds.]

---

an analysis would exploit the availability, in Indonesian, of zero-marked partitive expressions, as exemplified earlier in (5), in which *kaum* is understood as 'of the community'. In the present case, then, the numeral *dua* would be understood as meaning 'of a set of two'.

19 A possible way of extending the distributivity analysis to include intensification might involve considering the intensified adjective as distributing over a set of points on a scale. For example, in (37), *panjang-panjang* would be asserting that for each of a plurality of points on the scale of length (and by implicature perhaps a large number of such points), the host *panjang* bears the values of length associated with that point.

Examples (38)–(41) instantiate a variegated set of functions associated with reduplication. In none of them, however, is there any obvious plural set of entities that might be available as a potential distributive key licensing a relationship of distributivity. Accordingly, for the time being at least, they do not seem straightforwardly amenable to the unified analysis for reduplication proposed above.

To this point, all of the examples of reduplication discussed above were ones that could be readily associated with a unique etic function. However, this is precisely because they were adduced in order to support the independent existence of each of the proposed individual functions. In actual language use, though, it is equally, perhaps even more common to encounter utterances in which a single occurrence of reduplication appears to be associated with two or more of the above functions simultaneously. For present purposes, such utterances may be broadly divided into two main groups: those in which all of the associated functions are analyzable in terms of distributivity and concomitant plurality, and those in which only some of the associated functions are analyzable as such.[20]

The first of the two main groups, involving exclusively distributive-plural functions, is illustrated by examples (42)–(48) below. To begin with, the simple-plural example in (27) may be contrasted with the following:

(42) *Yang kepala-kepala itu*
    REL   DISTR~head   DEM:DEM.DIST
    'The one with all the heads.'
    [Speaker attempting to establish reference to a computer game with lots of heads of different shapes and colours.]

On the face of it, (42), like (27), expresses simple nominal plurality. However, as suggested by the context, the reduplicated form *kepala-kepala* goes beyond asserting that the heads are more than one in number; it also emphasizes their diversity and large number. In (42), reduplication may thus be analyzed as asserting, simultaneously, the distribution of *kepala* over pluralities of objects, kinds, and associated locations. Such additional shades of meaning are very common in reduplicative constructions marking nominal plurality.

Now contrast the examples of reduplication marking plurality of arguments with the following examples, in each of which reduplication appears to be simultaneously marking the plurality of two different arguments:

(43) *Makan-makan apa?*
    DISTR~eat   what
    'What are you all eating?'
    [Speaker on phone to friends having dinner.]

---

20 Of course, there is a third group of utterances, in which reduplication is associated with two or more functions none of which seem amenable to an analysis involving distributivity; for reasons of space, these are not considered here.

(44) *Saya kasi-kasi   dengan saudara*
   1SG DISTR~give with   sibling
   'I'll give you each some.'
   [Streetside shaman and medicine vendor, offering to distribute free samples
   to his onlookers.]

(45) *Ini   tentang-tentang apa ini?*
   DEM:DEM.PROX DISTR~about   what DEM:DEM.PROX
   'What's all that about?'
   [Speaker asking about the notes that I'm taking in my notebook.]

In (43), *makan-makan* marks the plurality of the agent, the addressee and his
friends, and also the patient, the different things being eaten, asked about by the
interrogative *apa*. In (44), *kasi-kasi* marks the plurality of the patient, the shaman's
wares, and also the benefactive, the addressees expressed with the kinship term
saudara. And in (45), *tentang-tentang* marks the plurality of the theme, the notes in
the notebook, and the object of the preposition, the various things that they are
about, again asked about by the interrogative *apa*.

   The following three examples show some of the ways in which reduplication
may mark plurality of arguments alongside some other kind of plurality:

(46) *Ini   untuk apa, yang hijau-hijau?*
   DEM:DEM.DIST for   what REL DISTR~green
   'What are those for, those green things all over the place?'
   [Speaker asking about plants spread out on roadside embankment to prevent
   erosion.]

(47) *Kalau si   Pai ambil-ambil-ambil-ambil aja*
   TOP PERS Pai DISTR~take   NEG.FOC
   'Pai just takes things.'
   [Speaker complaining about his friend's behaviour.]

(48) *Kita panggil ojek   satu-satu*
   1PL call   motorcycle.taxi DISTR~one
   'Let's call some motorcycle taxis, one for each of us.'
   [Group of five people walking down the street, speaker makes a suggestion.]

In (46), *hijau-hijau* marks the plurality of the theme, the plants referred to by the
demonstrative *ini,* and also their many locations. In (47), *ambil-ambil-ambil-ambil*
marks the plurality of the patient, the things that he takes, and also the multiplicity
of associated taking activities, that is to say, iterativity. And in (48), *satu-satu* marks

the plurality of the agent, expressed by pronoun *kita*, and also the several associated motorcycle-calling activities, or, once again, iterativity.

Thus, in examples (42)–(48), reduplication appears to be marking two (or more) entities simultaneously as plural distributive keys. Crucially, though, there is no vagueness or ambiguity with regard to the choice of plural entities; rather, the two (or more) plural entities are involved in a unique and clearly-defined intended interpretation. What seems to be happening, then, is that reduplication is selecting as distributive key a plurality of ordered sets consisting of two (or more) elements drawn from the respective plural arguments. For example, in (42), *kepala-kepala* would distribute over a plural set of ordered triplets each consisting of a head, a kind, and a location; in (43), *makan-makan* would distribute over a plural set of ordered pairs each consisting of an eater and his or her food; and in (48), *satu-satu* would distribute over a plural set of ordered pairs each consisting of a person and his associated motorcycle-calling activity. Accordingly, examples (42)–(48) may also be accounted in terms of the analysis proposed earlier for examples (27)–(35), with reduplication marking its host as a distributive share seeking out the presence of another entity which it marks as plural distributive-key.

In many cases, though, reduplication appears to combine a clearly distributive-plural function with some other function whose connection to distributivity and plurality is less obvious; some examples of such constructions are given in (49)–(54) below. Examples (49) and (50) illustrate the combination of a plural/distributive function with that of intensification:

(49) *Gatal-gatal sekali*
    DISTR~itch very
    'I'm itching all over.'
    [Speaker swimming, encountering a jellyfish.]

(50) *Orang Cina giginya   ompong   kalau minum kopi pa-panas*
    person China tooth:ASSOC snaggletooth TOP drink coffee DISTR~hot
    'Chinese people lose their teeth from drinking very hot coffee.'
    [Speaker warning me not to drink coffee that's too hot.]

In (49), *gatal-gatal* marks the mereological plurality of the speaker's body and also expresses intensification of itchiness, while in (50), *pa-panas* marks the iterativity of the drinking activity while also expressing intensification of heat. Examples (51)–(54) illustrate the combination of a distributive-plural function with those of atelicity, negative polarity, concessivity and deprecation, exemplified earlier in pure form in (38)–(41) above:

(51) *Kita ngopi-ngopi   aja*
    1PL DISTR~AG:coffee NEG.FOC
    'Let's just have coffee.'
    [Group arguing about what to do next.]

(52) *Tak ada nampak-nampak*
NEG exist DISTR~see
'I haven't seen him for some time.'
[Speaker talking about friend.]

(53) *Bawa sini aja basah-basah*
carry LOC:DEM:DEM.PROX NEG.FOC DISTR~wet
'Have them bring it here even though it's all wet.'
[Getting ready to check out of hotel, the laundry still isn't done; speaker makes suggestion.]

(54) *Ini-ini aja terus keluar*
DISTR~DEM:DEM.PROX REL continue go.out'
'It's just these that keep on coming out.'
[Playing a computer game, speaker is frustrated that the wrong pieces keep on coming.]

In (51), *ngopi-ngopi* marks the plurality of its agent expressed by the pronoun *kita*, plus also atelicity. In (52), *nampak-nampak* marks the plurality, or durativity, of the activities of seeing, and also negative polarity. In (53), *basa-basa* marks the plurality of its understood theme, the laundry, and also concessivity. And in (54), *ini-ini* marks the plurality of the theme, the game pieces, together with deprecation.

How might the examples of reduplication in (49)–(54) be analyzed? Under a splitting approach, reduplication would be considered as instantiating cumulative exponence, combining the expression of two different concepts in each case: distributivity plus intensification in (49) and (50), distributivity plus atelicity in (51), distributivity plus negative polarity in (52), distributivity plus concessivity in (53), and distributivity plus deprecation in (54). An obvious disadvantage of this approach is that it lacks in explanatory power, treating the observed semantic patterns as mere accidents. It thus fails to account for the recurrence of a similar range of meanings associated with reduplication across a wide range of languages, as observed by Moravcsik (1978) and others. Moreover, it fails to provide a principled reason why Riau Indonesian reduplication works this way but not Kupang Malay *dong*, or for that matter also English *-s*.[21] Alternatively, then, one might choose to pursue a lumping approach. Taking the lead from the unified analysis proposed for reduplication in examples such as (27)–(35) and (42)–(48), one might attempt to

---

21 English plural *-s* itself exhibits homophony, also marking possession and the 3rd person singular present simple on verbs, however, one cannot imagine an English utterance, in context, containing a single form *-s* associated simultaneously with plurality plus also one of its other functions. In contrast, utterances such as those in (49)–(54), with reduplication simultaneously marking plurality/distributivity and some other function, are a dime a dozen in Riau Indonesian.

extend it to some or all of the other cases of reduplication, accounting for them, too, in terms of a relationship of distributivity obtaining between the reduplicated expression, the distributive share, and some other understood or overtly expressed entity, the semantically plural distributive key. Prima facie, it is not obvious, though, how the notions of distributivity and plurality might be shown to be relevant to functions such as atelicity, negative polarity, concessivity and deprecation. For now, then, a complete and more explanatory account for the entire range of functions of reduplication in Riau Indonesian must be left for future research.

3 Agreement and the syntax of number

Given the strongly isolating nature of Indonesian, many of the phenomena discussed in the preceding section are either syntactic or else of an intermediate nature straddling the boundary between syntax and morphology. Thus, forms such as numeral classifiers (Section 2.3.1), pronouns entering into inclusory pronoun and associative plural constructions (Section 2.3.3), and the plural marker *pada* (Section 2.4.1) appear to exhibit properties of separate words occurring within constructions of a syntactic nature, while even reduplication (Section 2.4.2) seems on occasion to shade into a syntactic construction involving the repetition of complete words.

As an isolating language, Indonesian is also completely lacking in any kind of agreement. Nevertheless, in one variety, Papuan Malay, pronouns may enter into a construction that appears, prima facie, to exhibit some of the properties of number and person agreement. In most varieties of Indonesian, taking sentence (1) *Ayam makan* and replacing *ayam* 'chicken' with a pronoun such as 3rd person *dia* to yield *Dia makan* bears no grammatical consequences whatsoever: *Ayam makan* and *Dia makan* display the same grammatical structure. While for typological purposes, pronoun is a useful comparative concept, in Indonesian, for the most part, pronouns exhibit the same grammatical behaviour as most other words in the language, their only distinguishing characteristics, as suggested in Section 2.2 above, being semantic, pertaining to person and number features. Not so, however, in Papuan Malay. Pronouns in Papuan Malay exhibit several distinctive grammatical properties which set them apart from other expressions in the language, and in so doing also from their pronoun counterparts in other varieties of Indonesian. One of the consequences of these distinctive properties is to provide for a richer menu of strategies for the expression of number.

The first distinctive property of Papuan Malay pronouns is that the short monosyllabic pronominal forms, at least, may be bound to their hosts. Donohue and Sawaki (2007) provide some arguments to the effect that these forms are most appropriately transcribed as clitics; however, I have been unable to replicate the data on which their arguments are based. So far I have been able to find just one piece of evidence in support of their bound nature, involving naturalistic orthography.

Speakers sometimes write the pronouns joined on to their hosts, as in the following example:

(55) E BAPA SAMAU JUAL TAPI NANTI SA PAKE APA BAPA TONG PIGI MAKAN
  SUDA ALAPAR SEKALI TONG MAKAN WARUNG MANAKA
  E bapa sa mau jual tapi nanti   sa pake apa. Bapa tong pigi
  FILL father 1SG want sell but FUT.PROX 1SG use what father 1PL go
  makan suda, sa lapar sekali Tong makan warung   mana ka?
  eat PFCT 1SG hungry very. 1PL eat   food.stall which Q
  'I want to sell it but then what would I use. Let's go eat already, I'm really
  hungry. Which food stall should we eat at?'
  [SMS message, speaker talking about his mobile phone, and then changing
  the topic.]

In the above example, the combinations *sa mau* and *sa lapar* are written as single words.[22] However, writers are generally inconsistent in this respect; even in the above example, there are three cases of pronouns being written separately, *sa pake*, *tong pigi* and *tong makan*. Nevertheless, such examples of joined-on pronouns occur frequently in the naturalistic orthography of Papuan Malay, but rarely if at all in that of other varieties of Indonesian. To the extent that such pronouns are indeed bound, this provides motivation for considering the constructions in question as involving argument indexing.

  A second distinctive property of Papuan Malay pronouns is their high frequency in texts. As noted in Section 2.2 above, most varieties of Indonesian go to great lengths to avoid using pronouns. In contrast, Papuan Malay uses them with exuberance, with a frequency comparable to if not actually greater than in a language such as English. This can be illustrated through consideration of various formulaic or otherwise high-frequency expressions. Across Indonesia, people ask 'Where are you going?' as a common greeting: while in most places the 2nd person pronoun will never be used, e.g. Riau Indonesian *Mau ke mana?* (want to where), in Papuan Malay it is very common, e.g. *Ko mo ke?* (2SG want to). When one takes a picture of somebody and they want to see it right away, in most parts of Indonesia the request will consist of just the verb, e.g. Riau Indonesian *Tengok* (look), whereas in Papuan Malay it will most commonly contain the 1st person pronoun, e.g. *Sa liat* (1SG see). Thus, in examples such as these, the pronoun constitutes an additional strategy for the expression of number.

  A third, related property of Papuan Malay pronouns is their propensity for occurring repeatedly, as illustrated below:

---

22 In the latter case, the spelling *ALAPAR* also reflects the common omission of initial *s* in certain function words, *sa* often being pronounced as *a*.

(56) *Tmn dua ini dong dua dpat usir drii rmh jdiii dong dua harus tinggal dimna*
boleh kam tlong ksih tempat untuk dong dua tinggal kaaa
Teman dua ini   *dong dua dapat usir dari rumah, jadi*
friend two DEM:DEM.PROX 3PL two get   expel from house become
dong dua harus tinggal di mana boleh, kam tolong kasi tempat untuk
3PL   two must remain LOC where can   2PL help give place for
dong dua tinggal ka
3PL   two remain Q
'These two friends got kicked out of home, so where can they stay, please
find a place for them to stay.'
[Facebook status update, on behalf of a couple of friends.]

In the above example, a new topic *Teman dua ini* 'these two friends' is introduced, and is subsequently referred to three times by the pronominal expression *dong dua* 'the two of them'. Written in this way, it is not clear whether the speaker had in mind the dual pronoun *dendua,* or, as was actually written, the periphrastic expression *dong dua*. However, in either case, the three repetitions of the expression would be bizarre in other dialects of Indonesian.

The fourth and final distinctive property of Papuan Malay pronouns distinguishing them from pronouns in other varieties of Indonesian is their frequent occurrence with a coreferential conominal expression; in other varieties of Indonesian such constructions are possible but very rare. For example, in (56) above, the three occurrences of *dong dua* are preceded by the conominal expression *teman dua ini*. Examples such as these lead Donohue and Sawaki (2007) to claim that Papuan Malay has subject-verb agreement; however, these constructions differ from prototypical instances of agreement in at least the following three respects: First, it is not clear to what extent the would-be agreement marker is bound; indeed, in (56) at least it clearly a free form.[23] Secondly, the occurrence of the putative agreement marker is always optional: in no case is it obligatory.[24] Thirdly, whereas agreement markers are typically backgrounded, repeated occurrences of expressions such as *dong dua* in (56) serve to reinforce their prominence in the discourse. Thus, following terminology introduced by Haspelmath (2013), it would seem more insightful to characterize such constructions as involving an argument index cooccurring with a conominal expression.

Strings such as *teman dua ini dong dua dapat usir dari rumah* in (56) present an interesting analytical quandary. Under the analysis assumed in the preceding

---

23 Thus, for example, *dong dua* can stand on its own as a referring expression; alternatively, when occurring in front of its host, additional words may potentially intervene between it and its host.
24 Donohue and Sawaki (2007) provide some contexts in which they claim that the marker in question is obligatory; however, in similar contexts, constructions without the marker can be found in Kluge (2017) and also in my own data.

paragraph, *dong dua* is an argument index forming a constituent with its host *dapat usir dari rumah*. However, under an alternative analysis, *dong dua* would form a constituent with the preceding expression *teman dua ini*, providing an instance of the PN-marking construction discussed in Section 2.3.3 previously. In fact, evidence from the placement of adverbs suggests that both analyses may be appropriate. Example (57) below shows that *kelemarin* 'yesterday' may be inserted either after *dong dua*, as in (57a), thereby supporting the constituency of *teman dua ini dong dua* and its analysis as a PN-marked expression, or alternatively before *dong dua*, as in (57b), thereby supporting the constituency of *dong dua dapat usir dari rumah* and the analysis of *dong dua* as an argument index marking *dapat usir dari rumah*:

(57) a. *Teman dua ini   dong dua kelemarin dapat usir   dari*
          friend two DEM:DEM.PROX 3PL   two   yesterday get   expel from
          rumah
          house

   b. *Teman dua ini   kelemarin dong dua dapat usir   dari*
          friend two DEM:DEM.PROX yesterday 3PL   two get   expel from
          rumah
          house

   'These two friends got kicked out of their home yesterday.'

(In addition to the two analyses suggested above, the possibility should also be acknowledged of a third "flat" analysis whereby, in (56), *dong dua* forms a constituent with neither the preceding *teman dua ini* nor the following *dapat usir dari rumah*.)

A similar structural ambiguity is evinced also by the string *Empong tendua ada main wifi* in example (4) above. As shown in (58) below, *skarang* 'now' may be inserted either after *tendua*, as in (58a), supporting the analysis of *Empong tendua* as a PN-marked expression, or alternatively before *tendua*, as in (58b), supporting the analysis of *tendua* as an argument index hosted by *ada main wifi*:

(58) a. *Empong tendua skarang ada   main wifi*
          Empong 1DU   now   PROG play wifi

   b. *Empong   skarang tendua ada   main wifi*
          Empong now   1DU   PROG play wifi

   'Me and Empong are online now.'

However, example (58) differs from its predecessor in an important way. While in (57) *dong dua* is strictly coreferential with *teman dua ini*, in (58) *tendua* wholly contains *Empong* in its reference. Thus, while (58a) exemplifies inclusory pronominal or associative dual constructions in accordance with the discussion in Section 2.3.3, (58b) instantiates discontinuous variants of the same construction – a discontinu-

ous inclusory pronominal or a discontinuous associative dual. As a discontinuous associative dual, example (58b) thus bears a resemblance, albeit just a superficial etic one, to the discontinuous associative plurals with *pada* in Jakarta Indonesian, in examples (23) and (24) earlier.

While the above discussion treats PN marking and argument indexation as two distinct constructions, under a broader and more abstract view, unifying across nominal and verbal domains, PN marking may be subsumed under a more generalized notion of argument indexation, as represented in the following table:

Tab. 6: Generalized Argument Indexation in Papuan Malay.

	pron X	*X pron*
V	ko liat (2sɢ see) 'you see' agent/experiencer/theme	liat ko (see 2sɢ) 'see you' patient/stimulus
N	ko anjing (2sɢ dog) 'your dog' possessor	anjing ko (dog 2sɢ) 'you dog' essant

Table 6 above proposes a generalized notion of argument indexation in accordance with two independent dimensions. Rows classify argument indexation in accordance with the category of the host X, either V(erbal), exemplified by *liat* 'see', or N(ominal), exemplified by *anjing* 'dog'. Columns distinguish between the position of the pronoun, here exemplified by the 2nd person singular *ko*, either before or after its host X. The four cells thus represent four different subcases, or subconstructions, of a generalized argument indexation construction, distinguished by their thematic role assignments. Within each cell, the first line provides an example of the construction, while the second line specifies the thematic role typically assigned to the pronoun in that particular configuration. Whereas for verbal hosts, these thematic role assignments reflect basic SVO word order, for nominal hosts, the left-hand cell represents the bare prenominal possessor construction, while the right-hand cell represents the PN-marking construction discussed above. Note that Table 6 makes reference to the role of essant introduced in Section 2.4.2 earlier. Thus, for *anjing ko* above, *anjing* assigns the essant role to *ko*, thereby asserting that the referent of the 2nd person singular pronoun *is* a dog.

As suggested in Table 6, then, the cases of apparent subject-verb agreement exemplified in the present section are just a particular case of a generalized notion of argument indexation, one that includes also the PN-marking construction discussed in Section 2.3.3. In all four subcases represented in Table 6, the generalized argument index *ko* is associated with a common number-plus-person marking function, thereby providing a vivid instantiation of the relative pervasiveness of number marking in Papuan Malay, in comparison to other varieties of Indonesian.[25]

---

25 The argument indexation scheme represented in Table 6 for Papuan Malay differs even from other closely related eastern Malay varieties such as Kupang, Alor, Ambon, Ternate and Manado

4 Semantics and discourse

Many of the semantic and discourse features of particular constructions involving number were already discussed earlier, in conjunction with their grammatical properties, in Sections 2 and 3 above. Underlying these diverse features, though, is a noteworthy commonality: in contrast to many other languages, the overt expression of number in Indonesian is strongly associated with discourse prominence, in the sense of Boye and Harder (2012).

Simply put, number isn't expressed unless it matters. Imagine a situation where there's a barking noise outside the house. English, like many other languages, forces you to choose between *There's a dog out there* and *There are some dogs out there*, thereby committing yourself to either singular or plural interpretations, even though, in many actual situations, you may not know, or care, whether there is one dog or several. In contrast, in Indonesian, you can simply say *Ada anjing di luar* (exist dog LOC outside), with no commitment to the number of dogs; indeed, you would only mark the noun for number, as in, say *satu anjing* (one dog) or *anjing-anjing* (DISTR~dog) if it were really important, that is to say, if the distinction between one or more dogs were really expected to play a role in the subsequent discourse.

Most or all of the examples of overtly expressed number considered in this chapter are ones in which number does indeed matter; in such cases, the expression of number is thus associated with discourse prominence. To cite just a couple of examples: In (26), by using the plural word *pada*, the speaker is highlighting the extent of his hunger by likening it to a crowd of demonstrators spread across multiple locations in his stomach. Similarly, in (34), by using the reduplicated form *masuk-masukkan*, the speaker is reinforcing the universal-quantificational effect of *semua* 'all' and thereby amplifying his boast to the effect that he will pocket all the balls. Other cases, discussed earlier in this chapter, in which the expression of number is associated with discourse prominence, include the discourse persistence function of numeral classifiers in Section 2.3.1 and the foregrounding function of argument indexation in Section 3. Thus, in Indonesian, the expression of number is generally foregrounded, associated with discourse prominence; few if any cases present them-

---

Malay. First, as noted previously, in other eastern Malay varieties there is little or no evidence for the argument index being bound to its host. Secondly, none of the other eastern Malay varieties have bare prenominal possession; instead, the possessor and possessed are separated by an intervening possessive marker *punya*, often reduced, as exemplified in example (7) *Pak Besa pung jadwal dong* for Kupang Malay. (Note that this construction is also available as an alternative to the bare possessive construction in Papuan Malay, cf. example (13) *pace ko punya pulsa*.) Nor, as discussed previously, do they have the PN-marking construction illustrated in the fourth cell of the table. Thus, the simple, general and largely symmetric picture of argument indexation represented in Table 6, and its concomitant pervasive system of person and number marking, would appear to constitute a relatively recent innovation in Papuan Malay.

selves in which it is part of the backgrounded information, present for purely formal, grammatical reasons.

As argued by Boye and Harder (2012), discourse backgrounding is a characteristic feature of grammatical categories, including, though not limited to, that of number. Accordingly, inasmuch as the kinds of constructions considered in this chapter tend to be discourse prominent, it may be questioned to what extent they really constitute instances of grammatical number, as opposed to a more lexically and syntactically based category of quantification.

5 Conclusion

In Indonesian, the category of number presents a janus-faced picture of contrasting richness and peripherality. On the one hand, Indonesian possesses a wide range of strategies for the expression of number, many of which exhibit a substantial amount of variation from one dialect to another. On the other hand, number plays an extremely minor role in the basic grammatical organization of the language. Most of the phenomena described here are best thought of as optional embellishments: while one often finds instances of two or more markings of number cooccurring in mutual reinforcement, for example reduplication and *pada* in (18), or reduplication and the universal quantifier *semua* in (34), one equally readily encounters significant stretches of speech containing no expression of number whatsoever. In this respect, the paucity and peripherality of number marking in Indonesian reflects the general ground plans of the language, with its IMA structure, as well as its low grammatical morpheme density in the sense of Gil (2015).

The description of number in Indonesian presented in this chapter offers a number of lessons of a more general nature for the field of linguistics. As a case study in microvariation, the present study highlights the ways in which closely related dialects may differ substantially with respect to grammatical features, thereby providing a rich vein of comparative data. A central theme of the chapter is the need to distinguish between comparative concepts, relevant for cross-dialectal and cross-linguistic studies, and the descriptive categories that are actually valid for individual dialects and languages. Thus, in Section 2.3.3, we saw that in Papuan Malay, what seems like a handful of constructions expressing a range of familiar functions including singulatives, additive plurals, associative plurals, and inclusory pronouns, is actually more appropriately analyzed in terms of a single, more general construction type, that of PN-marked expressions. Similarly, in Section 2.4.2, we saw that reduplication in Riau Indonesian, apparently expressing a wide range of functions including plurality, large size, iterativity, durativity, plurality of arguments and others, is, in large part at least, more insightfully analyzed as deriving from a single unified function of distributivity. In both cases, as in others, the etic typological approach leads towards fragmentation, suggesting analyses involving polyfunction-

ality or even homophony, while the alternative emic language-specific approach points towards unification, in the form of analyses invoking macrofunctionality and a single *gesamtbedeutung*. Linguistics must find a way to bridge between these two alternative approaches.

Acknowledgements

Abbreviations

The interlinear glosses in this chapter make use of the following abbreviations:

1	1st person;
2	2nd person;
3	3rd person;
AG	agent orientation;
ASSOC	associative;
AUG	augmentative;
CLF	classifier;
DEM	demonstrative;
DIST	distal;
DISTR	distributive;
DU	dual;
EMPH	emphatic particle;
EP	end-point orientation;
EXCL	exclusive;
EXHRT	exhortative;
FAM	familiar;
FILL	filler;
FOC	focus;
FUT	future;
INCL	inclusive;
IRRIT	irritative particle;

LOC	locative;
MUT	mutation;
NEG	negation;
PAT	patient orientation;
PERS	personal article;
PFCT	perfect;
PL	plural;
POSS	possessive;
PROG	progressive;
PROX	proximal;
PST	past;
Q	question particle;
REL	relative;
SG	singular;
SUPERL	superlative;
TOP	topic;
UNCERT	uncertainty;

References

Adams, Nikki & Thomas J. Conners. Forthcoming. Imposters and their implications for third person feature specification. *Journal of Linguistics*.

Allan, Keith. 1980. Nouns and Countability. *Language* 56. 541–567.

Boye, Kasper & Peter Harder. 2012. A Usage-Based Theory of Grammatical Status and Grammaticalization. *Language* 88(1). 1–44.

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149.

Choe, Jae-Woong. 1987. Anti-Quantifiers and a Theory of Distributivity. PhD Dissertation, University of Massachusetts, Amherst.

Chung, Sandra. 2000. Reference to Kinds in Indonesian. *Natural Language Semantics* 8. 157–171.

Conners, Thomas J., John Bowden & David Gil. 2015. Valency Classes in Jakarta Indonesian. In Andrej Malchukov & Bernard Comrie (eds.), *Valency Classes in the World's Languages*, 941–986. Berlin/Boston: De Gruyter Mouton.

Conners, Thomas J., Claudia M. Brugman & Nikki B. Adams. 2016. Reference tracking and noncanonical referring expressions in Indonesian and other discourse pro-drop languages. *NUSA* 60. 59–88.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Dalrymple, Mary & Suriel Mofu. 2012. *Dusner*. Munich: Lincom Europa.

Daniel, Michael. 1999. Semantic Types of Associative Plurals, Data from Bagvalal. In Ekaterina Rakhilina & Yakov Testelets (ed.), *Tipologija i Teorija Jazyka, Ot Opisanija k Objasneniju*, [Typology and Linguistic Theory, From Description to Explanation] 362–370. Moscow: Jazyki Russkoj Kultury.

Daniel, Michael. 2000. Tipologija Associativnoj Mnozhestvennosti [The Typology of Associative Plurals], PhD dissertation, Moscow: Moscow State University.

Daniel, Michael. 2019. Associative Plural as Indexical Category. *Language Sciences*, https://doi.org/10.1016/ (last accessed on 15 January 2022).

Daniel, Michael & Edith Moravcsik. 2005. The Associative Plural. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 150–153. Oxford: Oxford University Press.

Donohue, Mark & Yusuf W. Sawaki. 2007. Papuan Malay pronominals: Forms and functions. *Oceanic Linguistics* 46(1). 253–276.

Fabri, Ray. 1993. *Kongruenz und die Grammatik des Maltesischen.* (Linguistische Arbeiten 292). Tübingen: Niemeyer.

Fortescue, Michael. 1984. *West Greenlandic.* London: Croom Helm.

Gil, David. 1982. Distributive Numerals. PhD Dissertation, UCLA, Los Angeles.

Gil, David. 1988. Georgian Reduplication and the Domain of Distributivity. *Linguistics* 26. 1039–1065.

Gil, David. 1992. Scopal Quantifiers: Some Universals of Lexical Effability. In Michel Keffer & Johan van der Auwera (eds.), *Meaning and Grammar. Cross-Linguistic Perspectives,* 303–345. Berlin/New York: Mouton de Gruyter.

Gil, David. 1994. The Structure of Riau Indonesian. *Nordic Journal of Linguistics* 17. 179–200.

Gil, David. 1996. Maltese 'Collective Nouns': A Typological Perspective. *Rivista di Linguistica* 8. 53–87.

Gil, David. 2002. The Prefixes *di-* and *N-* in Malay / Indonesian Dialects. In Fay Wouk & Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*, 241–283. Canberra: Pacific Linguistics.

Gil, David. 2003. English Goes Asian; Number and (In)definiteness in the Singlish Noun-Phrase. In Frans Plank (ed.), *Noun Phrase Structure in the Languages of Europe,* 467–514. (Empirical Approaches to Language Typology, Eurotyp 20.7). Berlin/New York: Mouton de Gruyter.

Gil, David. 2004. Riau Indonesian *sama*, Explorations in Macrofunctionality. In Martin Haspelmath (ed.), *Coordinating Constructions*, 371–424. (Typological Studies in Language 58). Amsterdam: John Benjamins.

Gil, David. 2005a. From Repetition to Reduplication in Riau Indonesian. In Bernhard Hurch (ed.), *Studies on Reduplication*, 31–64. (Empirical Approaches to Language Typology 28). Berlin/New York: Mouton de Gruyter.

Gil, David. 2005b. Word Order Without Syntactic Categories: How Riau Indonesian Does It. In Andrew Carnie, Heidi Harley, & Sheila A. Dooley (eds.), *Verb First: On the Syntax of Verb-Initial Languages*, 243–263. Amsterdam: John Benjamins.

Gil, David. 2009. Riau Indonesian: What Kind of a Language Is It?. *Kongres Internasional Masyarakat Linguistik Indonesia (KIMLI) 2009*, Masyarakat Linguistik Indonesia and Universitas Negeri Malang, Malang, 29–61.

Gil, David. 2013a. Monosyllabic Words, Foot Structure, Ludlings and Dialectal Variation in Malay/Indonesian. Paper presented at the Seventeenth International Symposium on Malay/Indonesian Linguistics (ISMIL17), Padang, 9 June 2013.

Gil, David. 2013b. Riau Indonesian: A Language without Nouns and Verbs. In Jan Rijkhoff & Eva van Lier (eds.), *Flexible Word Classes, Typological Studies of Underspecified Parts of Speech*, 89–130. Oxford: Oxford University Press.

Gil, David. 2015. The Mekong-Mamberamo Linguistic Area. In Nick J. Enfield & Bernard Comrie (eds.), *Languages of Mainland Southeast Asia. The State of the Art*, 266–355. (Pacific Linguistics). Berlin/Boston: De Gruyter Mouton.

Gil, David. 2017. Isolating-Monocategorial-Associational Language. In Henri Cohen & Claire Lefebvre (eds.), *Handbook of Categorization in Cognitive Science*, 471–510. 2nd edn. Amsterdam: Elsevier.

Gil, David. 2020. What Does It Mean to Be an Isolating Language? The Case of Riau Indonesian. In David Gil & Antoinette Schapper (eds.), *Austronesian Undressed: How and Why Languages Become Isolating*, 9–96. Amsterdam: John Benjamins.

Gil, David, Uri Tadmor, John Bowden & Bradley Taylor. 2015. *Data from the Jakarta Field Station, Department of Linguistics, Max Planck Institute for Evolutionary Anthropology*, 1999–2015. [https://archive.mpi.nl/islandora/object/lat%3A1839_00_0000_0000_0021_10DE_A] (last accessed 15 January 2022).

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 58–90. Cambridge: MIT Press.

Haspelmath, Martin. 2004. Coordinating Constructions. An Overview. In Martin Haspelmath (ed.), *Coordinating Constructions*, 3–39. (Typological Studies in Language 58). Amsterdam: John Benjamins.

Haspelmath, Martin. 2010. Comparative Concepts and Descriptive Categories in Cross-Linguistic Studies. *Language* 86. 663–687.

Haspelmath, Martin. 2013. Argument indexing: a conceptual framework for the syntactic status of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin/Boston: De Gruyter Mouton.

Helmbrecht, Johannes. 2005. Politeness Distinctions in Pronouns. In Martin Haspelmath, Matthew Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 186–189. Oxford: Oxford University Press.

Hidayat, Lanny. 2010. The Acquisition of Verb Argument Structure in Basilectal Jakarta Indonesian, PhD Dissertation, University of Delaware, Newark, DE, USA.

Hopper, Paul J. 1986. Some Discourse Functions of Classifiers in Malay. In Colette G. Craig (ed.), *Noun Classes and Categorization*, 309–325. Amsterdam: John Benjamins, Amsterdam.

Jacob, June & Charles E. Grimes. 2011. Aspect and directionality in Kupang Malay serial verb constructions: Calquing on the grammars of substrate languages. In Claire Lefebvre (ed.), *Creoles, their Substrates, and Language Typology*, 337–366. Amsterdam: John Benjamins.

Kluge, Angela. 2017. *A Grammar of Papuan Malay*. Berlin: Language Science Press.

Lander, Yury. 2003. Nominal Reduplication in Indonesian: Challenging the Theory of Grammatical Change. Paper presented at the Seventh International Symposium on Malay/Indonesian Linguistics (ISMIL7), Nijmegen, The Netherlands, 29 June 2003.

Lichtenberk, Frantisek. 2000. Inclusory Pronominals. *Oceanic Linguistics* 39. 1–32.

Link, Godehard. 1991. Quantity and Number. In Dietmar Zaefferer (ed.), *Semantic Universals and Universal Semantics*, 133–149. Berlin: Foris Press.

Maurer, Philippe & the APiCS Consortium. 2013. Nominal plural marker and third-person-plural pronoun. In: Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures*, 96–99. Oxford: Oxford University Press.

Mboek, S. J. 1984. *Struktur Bahasa Melayu Kupang*. Proyek Penelitian Bahasa dan Sastra Indonesia dan Daerah, Jakarta.

Moravcsik, Edith A. 1978. Reduplicative constructions. In Joseph Greenberg (ed.), *Universals of Human Languages*. Vol. 3, 297–334. Stanford: Stanford University Press.

Moravcsik, Edith A. 2003. A Semantic Analysis of Associative Plurals. *Studies in Language* 27. 469–503.

Quine, Willard van Orman. 1969. Ontological Relativity. In Willard van Orman Quine. *Ontological Relativity and Other Essays*, 26–68. New York: Columbia University Press.

Saragih, Chrisma F. 2012. The practical use of person reference in Papuan Malay, MA thesis Radboud University, Nijmegen.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Atlantic Highlands: Humanities Press.

Sneddon, James Neil. 2006. *Colloquial Jakartan Indonesian*. Canberra: Pacific Linguistics.

Steinhauer, Hein. 1983. Notes on the Malay of Kupang (Timor). In James T. Collins (ed.), *Studies in Malay dialects: Part II*, *NUSA* 1. 42–64.

Tadmor, Uri. 2009. Loanwords in Indonesian. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the World's Languages. A Comparative Handbook*, 686–716. Berlin/Boston: De Gruyter Mouton.

Tjung, Yassir Nasanius. 2006. The Formation of Relative Clauses in Jakarta Indonesian, A Subject-Object Asymmetry, PhD Dissertation, University of Delaware, Newark, DE, USA.

Wouk, Fay. 1989. The Impact of Discourse on Grammar: Verb Morphology in Spoken Jakarta Indonesia, PhD Dissertation, UCLA, Los Angeles.

Wouk, Fay. 1999. Dialect Contact and Koineization in Jakarta, Indonesia. *Language Sciences* 21. 61–86.

Michinori Shimoji

14 Number in Japonic Family

Abstract: This chapter gives an overview of the number systems of the Japonic language family, examining Standard Japanese (SJ), Japanese mainland dialects and Ryukyuan languages. For the latter two, our focus will be on the Shiiba dialect of Kyushu and on the Irabu dialect of Miyako Ryukyuan, drawing on the field data collected by the present author. Generally speaking, Japonic language varieties exhibit common features such as (a) the number marking system sensitive to animacy, (b) the relevance of the category 'address noun', a class of lexical nouns which can be used as terms of address, including proper names and elder kin terms, and (c) the system of classifiers according to which referents are counted. On the other hand, Japonic language varieties vary with respect to, for example, the presence or absence of clusivity distinction (which is only found in Ryukyuan), and the interplay between plural marking and individuation (e.g. definiteness); while pluralization in SJ is limited to the individuated (definite) nouns, this is not the case in Shiiba.

1 Overview

The Japonic language family comprises Japanese and Ryukyuan. Ryukyuan is a group of endangered languages spoken on an island chain called Ryukyu Archipelago, which lies between Japan's mainland and Taiwan. Ryukyuan itself can be divided into two major subgroups, Northern Ryukyuan and Southern Ryukyuan (Pellard 2015). Japanese and Ryukyuan are not mutually intelligible, nor are Northern and Southern Ryukyuan languages (Shimoji 2010).

Japonic languages and their sub-varieties are typologically not homogeneous, but they do share a number of basic clause- and phrase-structure properties, i.e. verb-final, modifier-head order and dependent marking. Generally speaking, Japonic languages have agglutinating morphology. Suffixes and enclitics abound while prefixes and proclitics are scarce. In addition to affixation, compounding and reduplication are also common in word formation. Most Japonic varieties display the nominative-accusative alignment pattern.

With regard to number, most Japonic languages have a dichotomic number system in which singular and plural are distinguished, while some Ryukyuan varieties, especially Amami (Northern Ryukyuan), have a trichotomic system where singular, plural and dual are distinguished for personal pronouns (Section 2.2.2). The distinction in clusivity for the first person non-singular, i.e. exclusive and inclusive, is also common in Ryukyuan, while it is totally absent in mainland Japanese varieties (Section 2.2.2).

Pronouns and lexical nouns differ in number marking in two major ways. First, number marking is obligatory in pronouns. Second, as noted above, in most dialects a marked number such as dual is found in pronouns but not in lexical nouns, and if there is dual in lexical nouns then there is dual in pronouns as well. The lexical category 'address noun', described in detail in Section 2, is relevant in describing number in many Japonic languages. Certain plural marking strategies are restricted to pronouns and address nouns but are not attested in the other lexical nouns, etc.

Plural marking may take different forms in most Japonic varieties, but no distinction is made between associative and additive plurals; rather, plural marking depends on the lexical class of the noun to which the marker is attached. For example, in the Irabu dialect of Miyako Ryukyuan (Shimoji 2017), the plural marker for address nouns (kin terms and proper names) is *-ta*, while the plural marker for all other animate (including human) nouns is *-mmi*. The form *zjunzi-ta*, i.e. the plural form of the proper name *zjunzi*, may be interpreted either as a group of people represented by *zjunzi* or, much less commonly, as more than one person named *zjunzi*, though the latter interpretation is pragmatically and conventionally hardly natural for native speakers.[1] Likewise, the human common noun *siitu* 'student' is pluralized with *-mmi*, and the resulting form *siitu-mmi* may either be interpreted as a group of people associated with a particular, highly individuated student or as more than one student. What determines the semantic interpretation in terms of associative vs. additive plurality is the degree of individuation of the noun being pluralized.

In what follows I will start each section with an overview of Standard Japanese (SJ), followed by a sketch of dialectal variation with respect to the feature discussed in the section. In doing so I will try to deal with as many features of typological interest as possible, taking up examples from mainland Japanese dialects and from Ryukyuan. In particular, many examples are from the Shiiba dialect of Miyazaki Prefecture and from the Irabu dialect of Miyako Ryukyuan, as I have much fieldwork experience in these two languages.

2 Pronominal, nominal, and verbal number

2.1 Generalities

Unlike well-known European languages, Japonic languages do not exhibit agreement between an argument and its governing predicate, or between the head noun

---

1 When citing examples, I will follow the transcriptions of the work cited, except for examples from Standard Japanese, for which the orthography of the present author (based on the phoneme inventory of SJ) is used.

and its modifier. Number marking in Japonic languages is realized chiefly by number affixation (more specifically, suffixation) to the noun in question. Another strategy is reduplication, but it is unproductive and subject to a number of peculiar phonological, lexical and semantic restrictions as compared to the regular affixation strategy (see Section 2.5 for detail).

In all known Japonic languages, number marking is obligatory for personal pronouns but not for lexical nouns. In SJ, for example, the bare stem form of the first person pronoun, *watasi*, is always interpreted as singular, but the form *watasi-tati*, which additionally carries the plural suffix *-tati*, must be taken to be plural. Thus, the bare stem form can be said to carry a zero affix which designates singular reference, showing a paradigmatic contrast between the singular (zero) and the plural (*-tati*). By contrast, the bare stem form of a lexical noun *kodomo* 'child' may be interpreted either way, as illustrated in (1). The explicit marking of plurality with *-tati*, as in (2), unambiguously leads to plural interpretation (see Section 4.1.2).

(1) *kodomo=ga ason-de i-ru.*
child=NOM play-SEQ PROG-NPST
'(There is) one child playing'
'(There are) children playing'

(2) *kodomo-tati=ga ason-de i-ru.*
Child-PL=NOM play-SEQ PROG-NPST
*'(There is) one child playing'
'(There are) children playing'

As briefly mentioned in the introduction to this chapter, animacy, specificity and definiteness play a key role in the description of number marking in both Japanese and Ryukyuan in two major ways. First, personal pronouns are obligatorily marked for number in all Japonic languages, whereas number marking is impossible for inanimate nouns in most of them. Second, as mentioned earlier, in most Japonic varieties different number marking morphemes may be used in accordance with the animacy/definiteness of the noun in question. In the Irabu dialect of Miyako Ryukyuan, for example, five different plural suffixes are used: *-ti* (first person), *-du* (second person), *-nukja* (third person/demonstrative), *-ta* (a limited set of human nouns including proper names, kin terms, etc.) and *-mmi* (for all others). Table 1 introduces a Japonic version of the Noun Phrase Hierarchy (Silverstein 1976), which is a convenient tool for describing many number-related features of Japonic languages. In the table, the obligatoriness of number and its semantic effect for each class of nouns is indicated using the data from SJ.

What has been added to the Noun Phrase Hierarchy is the category 'Address nouns'. Address nouns are lexical nouns which can be used as terms of address,

Tab. 1: Noun Phrase Hierarchy and number marking in SJ.

	Personal pronoun	Address noun	Human noun	Animal noun	Inanimate noun
PL Status	Obligatory		Optional	Optional	Impossible
PL Meaning	Associative		Associative/ Additive	Additive	N/A

including proper names (e.g. SJ *Hanako* 'Hanako'), elder kin terms (e.g. SJ *(o)toosan* 'father', SJ *(o)kaasan* 'mother', SJ *ozisan* 'uncle', *(o)niisan* 'elder brother', *(o)neesan* 'elder sister', etc.).[2]

These nouns constitute a special class, as they are the only class of nouns for which number marking is obligatory just as in the case of pronouns. For example, the elder kin term *otoosan* 'father' obligatorily carries the plural suffix *-tati* as long as it refers to plural entities, either in address (3) or in the reference function (4).

(3) [When calling to the speaker's father and his company]
 *otoosan{-tati/*Ø}=wa doko=ni ik-u=no?*
 father-PL=TOP where=DAT go-NPST=SFP
 'Where are you (lit. father and his company) going?'

(4) [When asking the speaker's mother where the father and his company are going]
 *otoosan{-tati/*Ø}=wa doko=ni ik-u=no?*
 father-PL=TOP where=DAT go-NPST=SFP
 'Where are they (lit. father and his company) going?'

In fact, address nouns are like pronouns in that address nouns can be (or must be) used in place of personal pronouns: in SJ and most other Japonic varieties, it is a social norm to avoid using a plain pronoun for elders to refer to the addressee, and to use appropriate address nouns instead, making the address nouns fulfil the function of personal pronouns. The generalization is then made that nominals with the personal-pronominal function (i.e. personal pronouns and address nouns) are

---

2 In SJ, kin terms have two forms for one referent, e.g. *titi* and *(o-)too-san* for 'father', with the latter carrying the politeness suffix *-san* and (optionally) the honorific prefix *o-*. The former is only used for reference while the latter may be used for both reference and a term of address. It is the latter form that is identified as an address noun: you cannot call your father with *titi* but need to use *(o)too-san* instead. In many other dialects like the Irabu dialect of Miyako Ryukyuan, the same form may be used both for reference and address, as in: *zza* 'father', *mma* 'mother', *buza* 'uncle', *aza* 'elder brother', *ani* 'elder sister', etc.

obligatorily marked for number in Japonic languages. The associative plural meaning is obligatorily marked for address nouns, as demonstrated in (3) and (4) above.

A limited set of upper social status names like *sensei* 'teacher', *syatyoo* 'company president', etc. are human common nouns, but may be integrated into the address noun category. In SJ, for example, a proper name like Yamada may be compounded with a social status name (e.g. *Yamada-sensei* 'Teacher Yamada', *Yamada-syatyoo* 'President Yamada', etc.), and such compounds function exactly like a plain proper name, i.e. as an address noun. The proper name stem of the compound may be omitted, as in *(Yamada) sensei*, *(Yamada) syatyoo*, etc., and the status name still functions as an address noun. As a result, the social status names in this stranded construction take on the function of the omitted address noun and behave like the address nouns in terms of plural marking. As illustrated in (5), if the speaker intends to use the noun *sensei* 'teacher' as an address noun, plural marking is obligatory and has associative reading, while it is not in (6) where *sensei* simply functions as a lexical human noun (and the presence or absence of plural marking is semantically motivated; see Section 4).

(5) [When talking about Yamada and his company]
    (Yamada) sensei{-tati/*Ø}=wa (minna) kaet-ta.
    (Yamada) teacher-PL=TOP   all   leave-PST
    'Teacher (Yamada) and his associates (all) left.'

(6) [When talking about teachers who gathered in a conference]
    sensei{-tati/Ø}=wa (minna) kaet-ta.
    teacher-PL=TOP   all   leave-PST
    'The teacher and his associates (all) left.'

As mentioned above, address nouns may have the address function (i.e. second person pronominal function) as in (3) or the reference function (third person pronominal function) as in (4). In a situation where an elderly person speaks to a small child (as when a father talks to his child), (s)he may use the address noun also for first person reference, as in (7) and (8). As illustrated in (8), when the address noun *otoosan* 'father' is used in self-address, the pluralization is obligatory if it functions for 'us'.

(7) *otoosan=to issyo=ni   kaimono ik-u?*
    father=COM together=DAT shopping go-NPST
    '(Are you going to) go shopping with me (lit. father)?'

(8) *otoosan{-tati/*Ø}=to issyo=ni   kaimono ik-u?*
    father{-PL/*Ø}=COM together=DAT shopping go-NPST
    '(Are you going to) go shopping with us (lit. father:PL)?'

2.2 Pronominal number

2.2.1 Standard Japanese

The pronominal paradigm of SJ is summarized in Table 2. Several comments are in order.

First, number distinction between singular and plural is obligatory for personal pronouns and demonstrative pronouns for human referents, but not for the other pronouns, i.e. the demonstrative pronouns used for non-human referents, the interrogative pronoun *dare* 'who' and the reflexive pronoun *zibun*. The interrogative *nani* 'what' cannot be pluralized. The interrogative pronoun *dare* usually does not take number marking, though the plural form *dare-tati* is not completely unacceptable for native speakers, and the informal form *doitu-ra* 'who the hell (PL)' is perfectly acceptable. One environment in which *dare-tati* sounds natural is one where the referent in question is discourse-linked, i.e. 'who' is more like 'which one', as in (9).

(9) [Context: three groups of students are told to decide which group gives their presentation first, and the professor returns to the classroom and asks]
 sate dare-tati=ga   saki=ni   yar-u=no?
 OK   who-PL=NOM first=DAT do-NPST=NLZ
 'OK, who (=which group)'s gonna do first?'

It may also be possible to analyze that *dare* 'who' here is an interrogative word in place of a proper name, i.e. the expression *dare-tati* may designate associative plurality. Second, there are a number of stem forms for personal pronouns (see Iwasaki 2013, Irwin and Zisk 2019). For example, the first-person referent is designated by *watasi* (neutral for formality and gender), *watakusi* (very formal and neutral for gender), *ore* (informal and typical of male speech), *atasi* (informal and typical of female speech), etc. There are also varying forms for second person reference: *anata* (neutral for formality and gender), *kimi* (formal and for the socially lower ranked addressee; neutral for gender), *omae* (informal and neutral for gender), etc. The above table only lists the more neutral forms for personal pronouns. Generally speaking, second person reference via a personal pronoun is not a usual choice for a socially higher-ranked or respectable addressee; as noted earlier in the introduction to Section 2, the social norm requires the speaker to use an appropriate alternative, i.e. an address noun.

Third, SJ has three plural markers, *-tati*, *-ra* and *-domo*. The first one is functionally unmarked. The suffix *-ra* indicates informality and/or inferiority (i.e. the referent is shown to have a lower status than the speaker), though the third person *kare*, which used to be a distal demonstrative pronoun, usually takes *-ra*, with no effect of informality. The suffix *domo* is much less common and its occurrence is highly restricted. On the one hand, *-domo* occurs with the humble form of the first person singular *watakusi* (*watakusi-domo*); on the other, it may occur with nouns of offen-

Tab. 2: Pronominal number in SJ.

	Referent	Singular	Plural
Personal pronoun	1st person	watasi-Ø	watasi-tati ware-ware[3]
	2nd person	anata-Ø	anata-tati
	3rd person	kare-Ø (for male)	kare-ra
		kanozjo-Ø (for female)	kanozjo-tati/kanozjo-ra
Demonstrative pronoun	3rd person (informal)	koitu-Ø 'this guy' (proximate) soitu-Ø 'that guy' (mesial) aitu-Ø 'that guy' (distal)	koitu-ra soitu-ra aitu-ra
	non-human	kore 'this thing' (proximate) sore 'that thing' (mesial) are 'that thing' (distal)	kore(-ra) sore(-ra) are(-ra)
Reflexive pronoun		zibun 'self'	(zibun-tati)
Interrogative pronoun	human non-human	dare 'who' nani 'what'	(dare-tati)

sive expressions like *baka-domo* 'stupid bastards', *inu-domo* 'dogs', etc. Both usages demonstrate that *-domo* functions to downgrade the social status of the referent.

2.2.2 Dialectal variation

Many Japanese dialects have several different forms for reference to the same person, and especially for second person reference. For example, the Shiiba dialect of Miyazaki Prefecture distinguishes five forms for second person, *ware* (neutral), *waga* (informal), *unu* (very informal and offensive), *wasama* (very formal) and *konata* (extremely formal), whereas it has only one form for first person reference, *ore*.

With respect to number, most Japonic languages have a pronominal number system similar to that of SJ, with a two-way number distinction (singular vs. plural) for all persons. However, unlike SJ, it is uncommon for a Japonic language to have a dedicated third person pronoun which corresponds to *kare* in SJ. Rather, it is common for demonstrative pronouns to be used for third person reference, instantiating Bhat's (1994) 'two-person' system. The Shiiba dialect of Miyazaki Japanese is an example (Table 3).

In Shiiba, the suffix *-domo* is generalized for all plural forms. The interrogative pronoun *dare* 'who' additionally has a plural form *dad-dare*, which is underlyingly

---

3 The form *ware-ware* is a reduplicated form of *ware*, which is an archaic first person singular form and is not used without reduplication in contemporary SJ.

Tab. 3: Pronominal number system in Shiiba.

	Referent		Singular	Plural
Personal pronoun	1st person		ore-Ø	ore-domo/on-domo
	2nd person	neutral	ware-Ø	ware-domo
		very formal	konata-Ø	konata-domo
		formal	wasama-Ø	wasama-domo
		informal	waga-Ø	waga-domo
		offensive	unu-Ø	un-domo
Demonstrative pronoun	3rd person		kore-Ø (proximate) sore-Ø (mesial) are-Ø (distal)	kore-domo sore-domo are-domo/ad-domo
	non-human		kore (proximate) sore (mesial) are (distal)	kore(-domo) sore(-domo) are(-domo)
Reflexive pronoun			menme 'self'	menme(-domo)
Interrogative pronoun	human non-human		dare 'who' nani 'what'	dare-domo/dad-dare

dare-dare, i.e. a reduplication of the stem *dare*, a very common strategy of deriving a plural form in Japonic languages. It is noted that there are many other Japonic languages which allow plural marking for 'who'. No detailed research exists, however, to examine whether plural marking for 'who' interrogatives is conditioned by discourse factors as observed in SJ (see (9) above).

The pronominal number system of many Ryukyuan dialects can be more radically different from that of SJ. Some Ryukyuan dialects have a three-way distinction in number (singular, plural and dual). The three-way distinction is very common in the Amami languages of Northern Ryukyuan. The source of the dual suffix in Amami is the numeral which designates **tari* 'two people', a cross-linguistically common source of the dual morpheme.

Tab. 4: Amami number system (Yuwan dialect; Niinaga 2014).

	1st person	2nd person		3rd person (demonstrative) proximate/mesial/distal
		plain	honorific	
Singular	wan-Ø	ura-Ø	nan-Ø	kuri-Ø/uri-Ø/ari-Ø
Plural	waa-kja	ura-kja	naa-kja	kut-taa/ut-taa/at-taa
Dual	wa-ttəə	ura-ttəə	na-ttəə	*

Tab. 5: Plural marking patterns of clusivity in Ryukyuan

	Type A	Type B	Type C	Type D	Type E
	Common	Common	Uncommon	Very rare	None
EXCL	Stem-PL	Stem-PL	Stem-PL$_{excl}$	Stem$_{excl}$(-PL)	Stem$_{excl}$
INCL	Stem$_{incl}$	Stem$_{incl}$-PL	Stem-PL$_{incl}$	Stem$_{incl}$-PL	Stem$_{incl}$-PL
Example	Irabu	Aragusuku	Kamikatetsu	Yonaguni	n/a
Singular	*ban*	*ban*	*wan*	*anu*	
EXCL	*ban-ti*	*ban-ta*	*wa-nnaa*	*banu(-nta)*	
INCL	*duu*	*duu-ta*	*wa-chaa*	*ba-nta*	

Most Southern Ryukyuan languages have a two-way distinction in the first person non-singular (i.e. inclusive vs. exclusive); cf. the clusivity in the Irabu dialect of Miyako Ryukyuan in the plural.

(10) *ban-ti=a   mmja par-a-t=too.*
     1-PL.EXCL=TOP now   leave-THM-INT=EMP
     'We (exclusive) are going to leave now.'

(11) *duu=ja   mmja par-a-di=ru?*
     1.PL.INCL=TOP now   leave-THM-INT=Q
     'Why not leave now?'

Plural marking morphology that expresses the clusivity distinction in Ryukyuan languages varies in two respects: (a) whether the first-person pronominal stem is identical in the exclusive and inclusive forms and (b) whether the plural suffix is identical in the exclusive and inclusive forms (Table 5).

With regard to (a) and (b), cross-dialectally common patterns (Types A and B) make use of a special inclusive stem. The exclusive stem is the same as that used in the singular (e.g. *ban* '1SG' in Irabu). The special inclusive stem does not carry a plural suffix (Type A) or additionally carries (in a redundant way) a plural suffix (Type B). In Type C, uncommon but attested, the same stem is used in both functions and the distinction is made by two different plural suffixes. In Type D, the special exclusive stem and the special inclusive stem are used with the same plural suffix attaching to both, a pattern which is only found in Yonaguni. Note that the plural suffix on the exclusive in Yonaguni is not obligatory, and it is this optional use that makes Yonaguni unique among Ryukyuan. Indeed, all other attested patterns (Types A to C) have obligatory plural suffixation for the exclusive category. Type E is logically possible but not attested in Ryukyuan.

The generalization can then be made from the overall patterning in Ryukyuan that if the inclusive is explicitly marked for plural by a plural suffix, then the exclusive must also be marked. In fact, the pattern observed in Yonaguni is transitional,

whereby the older system (Type E) is changing towards Type D. This change is predictable based on the generalization above, as the change proceeds in such a way to make the above-mentioned cross-dialectal generalization applicable to this language.

2.3 Nominal number

2.3.1 Standard Japanese

Plural marking for lexical nouns is with *-tati*. The other two plural suffixes, *-ra* and *-domo*, which are also employed in pronominal plural marking, may be used for nouns but the unmarked choice is *-tati*, which is neutral in terms of honorification, politeness and humbleness. The rest of this chapter focuses on *-tati* plurals, unless otherwise specified.

Morphologically speaking, plural marking is obligatory for address nouns just like for pronouns (Section 2.1). Plural marking is not obligatory for all other lexical nouns, even those which designate human referents. Thus, whereas the address nouns *taroo* 'Taro' (proper name) and *niisan* 'elder brother' (elder kin term) are obligatorily marked for plural in (12) and (13), the lexical noun *gakusei* 'student' may be either marked by *-tati*, as in (14a), or left unmarked, as in (14b).

(12) a. *taroo-tati=wa (minna) kaet-ta.*
Taro-PL=TOP (all) leave-PST
'Taro and his fellows (all) left.'

　 b. **taroo=wa (minna) kaet-ta.*
Taro=TOP (all) leave-PST
'Intended meaning: Taro and his fellows all left.'

(13) a. *niisan-tati=wa (minna) kaet-ta.*
elder.brother-PL=TOP (all) leave-PST
'My elder brother and his fellows (all) left.'

　 b. **niisan=wa (minna) kaet-ta.*
elder.brother=TOP (all) leave-PST
'Intended meaning: My elder brother and his fellows all left.'

(14) a. *gakusei-tati=wa (minna) kaet-ta.*
student-PL=TOP (all) leave-PST
'The students (all) left.'

　 b. *gakusei=wa (minna) kaet-ta.*
student=TOP (all) leave-PST
'Student(s) (all) left.'

Semantically speaking, there is a difference between the *-tati* plural (14a) and the bare noun (14b), as the translations indicate. See Section 4.1.2 for more detail. In this sense, it is misleading to conclude that plural marking in SJ is 'optional'.

2.3.2 Dialectal variation

All known Japonic languages have a two-way number distinction for lexical nouns, singular vs. plural, one potential exception being the Yoron dialect of Amami Ryukyuan (Yamada 1985), which is claimed to have the dual category for address nouns (as in *uja-tai* 'parent-DU') as well as for pronouns (Yamada 1985).

The majority of Japonic languages have one plural suffix generalized for both pronouns and lexical nouns, exhibiting the *generalizing* pattern, while some dialects exhibit the *differential PL marking* pattern, where the choice of the form of the plural suffix is largely conditioned by the animacy-definiteness of the noun to which it is attached.

The Shiiba dialect of Miyazaki Prefecture exhibits the generalizing pattern, with the same plural suffix *-domo* for all lexical nouns as well as for pronouns, as in *ore-domo* (1st person-PL; see Section 2.2.2 for other pronouns), *taroo-domo* 'Taro-PL', *togi-domo* 'friend-PL', *inu-domo* 'dog-PL', etc.

The plural marking of the Miyoshi dialect of Chiba Prefecture (Sasaki 2018) is an example of the differential pattern, where the suffix *-dara* is restricted to the first and second person pronouns; the suffix *-ra* marks the third person demonstrative pronouns, address nouns and human common nouns; and the suffix *-domo* only occurs with human common nouns.

The Irabu dialect of Miyako Ryukyuan exhibits the richest differential PL marking pattern among all Japonic languages. Each pronoun takes a unique plural suffix, *-ti* for the first person pronoun (exclusive), *-du* for the second person pronoun, *-nukja* for the third person demonstrative pronoun; for lexical nouns, *-ta* is used for address nouns (i.e. proper names and elder kin terms) as in (15) and *-mmi* for all other animate nouns as in (17). Some nouns, specifically the social status names (see Section 2.1), take either *-ta* or *-mmi*, as in *sinsii{-ta/-mmi}* 'teacher:PL', depending on whether the term functions as an address noun (the title without a proper name; (16a)) or as a simple noun ('teacher'; (16b)).

(15) *anna-ta=a*   *(mmna) par-i-i*   *njaa-n.*
    mother-PL=TOP (all)   leave-THM-SEQ PRF-NPST
    '(The) mothers all left' [additive reading]
    'The group of people represented by a mother (all) left.' [associative reading]

(16) a. (Yamada) *sinsii-ta=a*   *(mmna) par-i-i*   *njaa-n.*
      (Yamada) teacher-PL=TOP (all)   leave-THM-SEQ PRF-NPST
      'Teacher (Yamada) and his fellows (all) left' [associative reading]

b. *sinsii-mmi=a*   *(mmna) par-i-i*   *njaa-n.*
   teacher-PL=TOP (all)   leave-THM-SEQ PRF-NPST
   'The teachers all left' [additive reading]

(17) *siitu-mmi=a*   *(mmna) par-i-i*   *njaa-n.*
    student-PL=TOP (all)   leave-THM-SEQ PRF-NPST
    '(The) students all left' [additive reading]
    'The group of people represented by a student all left.' [associative reading]

It is a common feature of the Japonic family for plural marking to be restricted to animate nouns, especially human nouns and less commonly animals like dogs, cats, horses, deer, etc., which are closer to humans in the sense that they are domesticated and/or salient in some other way in the human culture of the language in question. The Shiiba dialect of Miyazaki Prefecture, for example, allows plural marking on animal nouns like *inu* 'dog' and *kantyoo* 'deer', as in *inu-domo* 'dogs' and *kantyoo-domo* 'deer:PL' (probably due to the fact that they are culturally more salient than other animals in the hunting society), while it never allows plural marking on other animate nouns like *arimusi* 'ant' and *io* 'fish'.

    There are some dialects which do mark plurality for inanimates, as in Kansai dialects (Kanbayashi 2017, Oshima 2018) and the Kudaka dialect of Okinawa Ryukyuan (Niinaga 2015). These languages tend to show stronger restrictions on the plural marking for inanimates than for animates. For example, in the Kudaka dialect (Niinaga 2015), all the examples of the plural marking on inanimate nouns listed by Niinaga additionally carry the diminutive suffix *-gwaa*, as in *hii-gwaa-taa* (tree-DIM-PL) 'twigs', *hadibaa-gwaa-taa* 'little dead leaves'. The effect of the diminutive marking must relate to increasing animacy (or empathy), which might perhaps explain why inanimate nouns start taking plural marking. For other dialects which may allow plural marking for inanimate nouns, it tends to mean approximation rather than genuine plurality; this use of the plural marking is absent in SJ (see Section 4.3 for a discussion of the semantics of approximation).

2.4 Verbal number

There is no 'verbal number' as such in Japonic varieties including SJ and Irabu Ryukyuan. Some scholars claim that Classical Japanese had a converbal form which functioned to indicate plurality of its subject (Kazama 2003), as illustrated in (18).

(18) *mina koki*   *koromo=o ki-te*   *hikikaesi-tutu ari-si=ni* ...
    all   deep.colored clothes=ACC wear-CVB turn.down-CVB RSL-PST=DAT
    '(They) all wore deep colored clothes, turning the collars down ...'
    (Hayashi and Ando 1997: 886, cited in Kazama 2003: 255; Makurano Soshi)

Referring to Hayashi and Ando (1997), Kazama (2003) notes the plural marking function of *-tutu*. Hayashi and Ando (1997) state that the converb suffix *-tutu* "indicates simultaneity of actions performed by more than one agent". Even though there is no research which addresses the issue of whether Japanese (or Classical Japanese) has verbal number as such, it is noteworthy that if the above-mentioned analysis is correct, this morpheme appears to mark verbal number. It is not an aspect marker stricto sensu, because it obligatorily indicates the plurality of the agents concerned.

2.5 Reduplication as number marking

Reduplication as a morphological means is quite common in all Japonic languages including SJ, and it has various functions such as iterativity, distributivity, etc. Reduplication may be used for pluractionality but does not necessarily indicate or imply the existence of plural referents. For example, in the Irabu dialect of Miyako Ryukyuan, *mii-mii* 'look~RED' may express a situation where a single agent is looking at a single object again and again, demonstrating that the effect of reduplication is that of iterativity rather than pluralization of the arguments involved.

In SJ and many other mainland Japanese dialects, reduplication of a nominal stem also has the function of marking nominal plurality. In SJ, pluralizing reduplication shows a number of peculiarities and is subject to many more restrictions than the affixation strategy. First and foremost, it is highly lexicalized and unproductive. For example, *hana* 'flower', *ki* 'tree' and *eda* 'branch' may be pluralized by reduplication, as in *hana-bana, ki-gi* and *eda-eda* respectively, but *na* 'vegetable', *mi* 'nut', *ne* 'root', *kuki* 'stem', etc., cannot[4]. For the first person, its plurality may be expressed by *ware-ware* '1PL (formal)', but the corresponding singular stem *ware* is archaic and is not used in contemporary SJ (see footnote 3). As Tamamura (1985) points out, there are many other phonological, lexical and semantic restrictions with regard to the usage of nominal reduplication in SJ. For example, there are many reduplicated words with one or two morae as noted above, but it is very difficult to find examples of words with three or more morae. With respect to semantics, the reduplication strategy often expresses heterogeneity of the plural group rather than simple plurality. Thus, *kami-gami* 'gods', which is a reduplicated form of the noun stem *kami* 'god', refers to a group of highly individuated gods of different kinds, rather than a simple homogeneous group of gods. Unlike the affixation strategy, which is more likely to be used for animate referents, the reduplicated strategy

---

4 The initial consonant of the reduplicant in each example may undergo "sequential voicing", whereby a voiceless obstruent is replaced by the voiced counterpart with the same point of articulation (/k/ → /g/; /h/ exceptionally alternates with /b/ as the former used to be */p/ in proto-Japonic). Sequential voicing is typical of compounding, and we can see that the reduplicant in a reduplicated word behaves like a compound stem in this regard.

may be used for inanimate ones, especially spatio-temporal notions like *jama-jama* 'mountains', *hi-bi* 'days'.

Generally speaking, Ryukyuan languages lack pluralizing reduplication or use it in a much more limited way than mainland Japanese varieties. In Northern Ryukyuan, especially in Amami Ryukyuan, the reduplication strategy seems to be more prominent than in Southern Ryukyuan. In the Okinoerabu dialect of Amami Ryukyuan (Yokoyama 2017), for example, the interrogative pronoun *taru* 'who' may be reduplicated to indicate plurality, as in *taruu-daru* 'who:PL'. In the Tokunoshima dialect of Amami Ryukyuan (Masahito Ideguchi, p.c.), the reflexive pronoun *duu* 'oneself' may be pluralized by reduplication, as in *duu-duu* 'one selves'.

3 Agreement and the syntax of number (with special focus on numerals and classifiers)

3.1 Standard Japanese

Japonic languages lack agreement between the head noun and its modifier or between the verb and its arguments.

SJ is characterized by its rich inventory of numerals and classifiers (see, for example, Martin 1975, Iwasaki 2013 and Irwin and Zisk 2019 for fuller accounts of the classifier-numeral system). There are three sets of numerals, each of which belong to different lexical strata.

(19)   Numerals in SJ
   a. Native: *hi*(*to*)- (1), *hu*(*ta*)- (2), *mi-* (3), *jo-* (4), *itu-* (5), *mu-* (6), *nana-* (7), *ja-* (8), *kokono-* (9), *too* (10)
   b. Sino-Japanese: *iti* (1), *ni* (2), *san* (3), *si* (4), *go* (5), *roku* (6), *siti* (7), *hati* (8), *kjuu* (9), *zjuu* (10)
   c. English: *wan* (1), *tuu* (2), *surii* (3), *foo* (4), *faibu* (5), *sikkusu* (6), *sebun* (7), *eito* (8), *nain* (9), *ten* (10)

The last series is restricted in use, and is mostly used in games using the terminologies adopted from English, such as *wan-auto* 'one out (used in baseball)', *tuu-auto* 'two outs', etc. For counting numbers, Sino-Japanese or native bare numerals are used, but the latter are going out of use; *hii, huu, mii, joo*, etc., are not usually used, and the Sino-Japanese series, *iti, ni, san, si*, etc., are used, though those of the native series may also be used, in which case monomoraic stems are lengthened, as in *hii* 'one', *huu* 'two', *mii* 'three', *joo* 'four', etc. For counting objects, these numeral roots must be combined with an appropriate classifier suffix according to the category of the referents being counted. For example, sharp and long objects like pencils, um-

brellas and trees constitute a class which are counted with the classifier suffix *-hon*, while sheets of paper and cloth constitute another class and are counted with the classifier *-mai*; etc. A structure like that in (22), with no classifier suffix, is ungrammatical.

(20) *ni-hon=no   enpitu*
     2-CLF=GEN pencil
     'two pencils'

(21) *ni-mai=no   kami*
     2-CLF=GEN paper
     'two sheets of paper'

(22) **ni=no enpitu/kami*
     2=GEN pencil/paper

The classifier system of SJ is sensitive to various features of the referent, including, among others, (a) animacy, (b) physical properties, and (c) properties of events. The following list is a non-exhaustive list of major classifier suffixes found in SJ.

(23)   Major classifier suffixes in SJ
       a. human: *-ri*/*-nin* (*-ri* for the lowest numerals, i.e. *hito-ri* 'one person' and
          huta-ri 'two persons; and *-nin* for all others, e.g. *san-nin* 'three persons',
          yo-nin 'four persons')
       b. animal: *-too* (specifically for relatively big animals like elephants, whales,
          etc., as in *san-too=no zoo* 'three elephants'), *-wa* (specifically for birds), *-hiki*
          (for animals of other kinds or in general)
       c. object: *-tu* (the most general classifier for counting objects), *-ko* (usually
          used to count small objects), *-satu* (used to count books), *-mai* (used to
          count papers and other thin and flat objects like notes, dishes, etc.), *-hon*
          (used to count small and sharp objects like pens and needles), *-bi* (used
          specifically to count uncooked fish), etc.
       d. event: *-kai* (the most general classifier for counting events), *-hatu* (for count-
          ing punctual events like explosions, punches, etc.), *-niti* (for counting days),
          -zikan (for counting hours), etc.

A numeral with a classifier may precede the noun bearing the genitive case marker as a modifier of the NP headed by the noun, as in (20) and (21) above, or may 'float' to outside the NP, as in (24), stripped of case marking. Furthermore, as an independent word, a numeral with a classifier may stand as a head of an NP, as in (25).

(24) *enpitu=ga   ni-hon aru.*
    pencil=NOM 2-CLF   exist
    'There are two pencils.' (cf. (1))

(25) *ni-hon=o   koko=ni,   moo   ni-hon=o   asoko=ni   oite.*
    2-CLF=ACC here=DAT another 2-CLF=ACC there=DAT put.IMP
    'Put the two (of the pencils) here, and put another two there.'

3.2 Dialectal variation

All known Japonic dialects have comparable classifier systems, with a wide range of classifiers which are sensitive to the physical properties, animacy or other features of the noun being counted. However, there are at least two conspicuous dialectal discrepancies between SJ and (at least some) Japonic dialects.

First, most dialects have a much smaller inventory of numerals and classifiers than SJ due to a lack of classifiers of Sino-Japanese and/or English origin. For example, whereas SJ uses different classifiers for big animals (*-too*) and small animals (*-hiki*), the Irabu dialect of Miyako Ryukyuan uses a single classifier *-kara* for both bigger and smaller animals. Also, Irabu only has a native numeral set, which is cognate to the native numeral set in SJ, which goes back to proto-Japanese-Ryukyuan.

A second feature of dialectal variation concerns the order of the numeral and the noun. In SJ, both the prenominal and the postnominal numeral are allowed (e.g. (20) vs. (24) above), while some dialects rarely allow the prenominal numeral. The Yuwan dialect of Amami Ryukyuan (Niinaga 2014) is such a language. In Yuwan, while prenominal modification by a numeral is possible, as in (26), the most common way of using a numeral word is to make it the head of an NP, as in (27), where the noun stands as a modifier of the NP and carries the genitive case *=nu* (Niinaga 2014: 238–239).

(26) *un   k²wanu, c²jui=nu   k²wa=nu   isjoobiki hucji,*
    that child   one.CLF[HUMAN]=GEN child=NOM whistle blow.SEQ
    (24) 'That child…one of the children whistled, and …'
    [prenominal modification]

(27) *un   kago=nu   t²ɨɨ   cɨdɨ   ikjun=wake.*
    that basket=GEN one.CLF[THING] load.SEQ go=DSC
    '(The boy) puts that one of the baskets on (the front of his bicycle) and goes.'

As Niinaga states, the structure exemplified by (27) differs from the so-called 'quantifier float' analysis, which is common in analyzing post-nominal modification by a

numeral in SJ (24). In a quantifier float construction, the head of an NP is the noun being counted, and the underlying modifier of the NP, i.e. the numeral, floats outside of the NP. This is obviously not the case in (27), where the noun being counted carries the genitive case, indicating its dependent status. Yuwan does have a quantifier float construction, as shown in (28), where the numeral *cʔjui* occurs outside of the NP, bounded by the nominative case marker *=ga*.

(28) *un=nənti uzii=ga   cʔjui   joonasi mutun=wake.*
      hat=LOC old.man=NOM one.CLF[HUMAN] pear   be.picking.up=DSC
      'There is an old man picking up a pear there.' [floating]

4 Semantics and Discourse

4.1 Standard Japanese

4.1.1 Bare noun vs. *-tati* plural

As mentioned in Section 2.3.1, plural marking for non-pronominal nouns (i.e. lexical nouns excluding address nouns) is not morphologically obligatory; cf. (29b).

(29) a. *gakusei-tati=wa minna kaet-ta.*
         student-PL=TOP all   leave-PST
         'The students all left.'

     b. *gakusei=wa   minna kaet-ta.*
        student=TOP all   leave-PST
        'Student(s) all left.'

However, this does not mean that plural marking is simply optional. Rather, a *-tati* plural and a bare noun have different semantic interpretations when used in a plural context. Many scholars thus point out that *-tati* plurals in SJ exhibit properties pertaining to a broad concept of pragmatic salience (animacy, definiteness and specificity). That is, a *-tati* plural makes the pluralized set of referents pragmatically salient (definite and/or specific). Some scholars ascribe the behaviors of *-tati* specifically to its definiteness (Kurafuji 1999, 2003) while others suggest alternative analyses with reference to the potential non-uniformity of the *-tati* pluralization (Nakanishi and Tomioka 2004) or to specificity (Hosoi 2005, etc.). Interested readers are referred to these works for more detail on the semantics of *-tati*.

The pragmatic differences between bare vs. *-tati* plurals are demonstrated in the following three grammatical contexts. First, a bare noun may have a generic/ kind interpretation, while a *-tati* plural cannot (Nakanishi and Tomioka 2004).

(30) a. *Itariazin=wa yooki=da.*
    Italian=TOP cheerful=COP
    'Italians are cheerful.'

   b. *Itariazin-tati=wa yooki=da.*
    Italian-PL=TOP   cheerful=COP
    ??'Italians are cheerful.'
    'The Italians are cheerful.'

Second, while a bare noun can be used as the internal argument of a relational possessive expression like 'I have kids', a *-tati* plural cannot.

(31) a. *watasi=ni=wa kodomo=ga i-ru.*
    1SG=DAT=TOP child=NOM exist-NPST
    'I have a child/children.'

   b. **watasi=ni=wa kodomo-tati=ga i-ru.*
     1SG=DAT=TOP child-PL=NOM exist-NPST
    'I have a child/children.'[5]

Third, a bare noun takes a narrow scope in intentional predicates like 'seek for', while a *-tati* plural takes a wide scope (examples taken from Nakanishi and Tomioka 2004: 115).

(32) a. *sono byooin=wa   kangohu=o sagasite   i-ru.*
    that hospital=TOP nurse=ACC look.for.GER PROG-NPST
    'This hospital is looking for a nurse/nurses (to hire).'

   b. *sono byooin=wa   kangohu-tati=o sagasite   i-ru.*
    that hospital=TOP nurse-PL=NOM look.for.GER PROG-NPST
    *? 'This hospital is looking for a nurse/nurses (to hire).'
    'There is a group of nurses that hospital is looking for.'

The above-mentioned individuation-related features of plural marking recur in the plural marking of other languages such as Mandarin Chinese (Iljic 1994), Korean (Kwon and Zribi-Hertz 2004), Malay (Nomoto 2013) and Persian (Ghomeshi 2003, Nomoto 2013). In a formal theoretical perspective, Nomoto (2013) hypothesizes that these languages have a pure quantifier system in common, in contrast to a determiner quantifier system like that in English.

---

5 (31b) is ungrammatical if the intended meaning is a description of the speaker's parenthood as in (31a), but may sound natural in a situation where the speaker has lost almost everything except his/her children, e.g. after a terrible earthquake destroying their house. Here, the plural-marked "children" is definite.

4.1.2 Associative vs. additive plural

SJ makes use of the single marker *-tati* for associative and additive plurals, exhibiting a cross-linguistically very common pattern. The associative vs. additive reading is heavily influenced by the degree of pragmatic salience of the noun to which *-tati* is attached, reflecting a cross-linguistically common tendency for the associative reading to appear at the higher end of the Animacy Hierarchy (Corbett 2000: 86).

(33) Personal pronoun: *watasi-tati* (1st person-PL): 'a group represented by me' (associative PL)

(34) Address noun:
a. Proper name: *taroo-tati* (Taro-PL): 'Taro and his company' (associative PL)
b. Elder kin term: *neetyan-tati* (elder sister-PL)
     'The elder sister and her company' (associative PL)
     'The elder sisters' (additive PL)

(35) Human common noun: *otoko-tati* (man-PL):
     'the man and his company' (associative PL)
     '(the) men' (additive PL)

Associative interpretation is impossible with a bare noun. Let us examine the following pair from Nakanishi and Tomioka (2004: 128). (36a) and (36b) differ only with respect to the plural marking with *-tati* on *onnanoko* 'girl'. In (36a), the *-tati* form can refer to a group which includes some entities other than girl(s) (in this example, boys) because of the associative nature of the *-tati* suffix. The bare form, by contrast, always yields an additive reading and cannot refer to entities other than girl(s) (36b).

(36) a. *kooen=de utat-te   i-ta   onnanoko-tati=no naka=ni=wa*
        park=at   sing-SEQ PROG-PST girl-PL=GEN   inside=DAT=TOP
        otokonoko=mo ni-san-nin mazat-te   i-ta.
        boy=also   2–3-CLF   be.included-SEQ PROG-PST
        'Among (the) girls who were singing in the park, a few boys were included.'

   b. *??kooen=de utat-te   i-ta   onnanoko=no naka=ni=wa*
        park=at   sing-SEQ PROG-PST girl=GEN   inside=DAT=TOP
        otokonoko=mo ni-san-nin mazat-te   i-ta.
        boy=also   2–3-CLF   be.included-SEQ PROG-PST
        'Among (the) girls who were singing in the park, a few boys were included.'

4.2 Dialectal variation

4.2.1 Bare noun vs. explicit plural marking

As in the case of SJ, plural marking is morphologically optional for lexical nouns in all known Japonic dialects. Due to the lack of detailed research on the plural forms and their dialectal variation, it is unclear whether the same contrast between the bare noun and the explicit plural marking as that in SJ is found in Japonic dialects. The author has the relevant data from the Shiiba dialect of Miyazaki Prefecture, which shows some important divergence from the situation in SJ.

First, the explicit plural marking with *-domo* is compatible with a generic interpretation. Compare (37) from Shiiba with (30) above from SJ. Though the exact semantic difference between the *-domo* plural and the bare noun is still unknown, native speakers report that the two are interchangeable in a generic context, a judgment which never obtains for SJ (30).

(37) *itariazin(-doma)=a yookina=nee.*
    Italian-PL=TOP   cheerful=SFP
    'Italians are cheerful.' (Note: domo + =a → *domaa* by a phonological rule)

Second, unlike SJ, the internal argument of a relational possessive expression like 'I have kids' can be marked with *-domo*. Again, *-domo* is reported to be optional and there is no known semantic difference between the *-domo* plural and the bare noun.

(38) *ore=nyaa   kodomo(-domo)=no (san-nin)   or-u.*
    1SG=DAT.TOP child(-PL)=NOM   three-CLF[HUMAN] exist-NPST
    'I have (three) kids.' (intended as a description of the speaker's motherhood)

Third, unlike SJ, a bare noun as well as a *-domo* plural may take a narrow scope in intentional predicates like 'seek for'.

(39) *akko=no   koomuten=wa   syokunin(-domo)=o*
    that.place=GEN building.contractor =TOP carpenter(-PL)=ACC
    sagas-i-t-or-u=tyuu=wai.
    seek.for-THM-IPFV-PROG-NPST=HS=SFP
    'That building contractor's company is said to look for carpenters for hire.'

All these examples indicate that plural markers in Japonic dialects may have a wide range of dialectal variation with regard to semantics. Some dialects, like SJ, have a plural marker whose function has much to do with pragmatic salience and there is a strong restriction on its use in generic contexts, while others, like the Shiiba dia-

lect, have a plural marker whose function is not apparently connected to the pragmatic salience.

4.2.2 Associative vs. additive plural

Like SJ, most Japonic dialects use the same plural marker to encode both associative and additive plural, depending on the animacy-definiteness of the noun to which it is affixed. In the Shiiba dialect of Miyazaki Prefecture, *-domo* denotes associative plural when attached to a personal pronoun (*ore-domo* 'we'), but it may denote either associative or additive plural when attached to a human common noun, as in *otoko-domo* ('the man and his company' or 'men').

Some dialects have a disambiguation strategy which pertains to the associative-additive interpretation. The Yoron dialect of Amami Ryukyuan (Kibe et al. 2019) has a special tone shift which unambiguously designates additive plurality. In the following pair of examples, the symbol '[' indicates where the level high tone begins.

(40) a. *[ʔatɕa-taa* 'father-PL' 'a set of people associated with a particular father'
    b. *ʔatɕa-[taa* 'father-PL' 'fathers'

In (40a), the level high tone begins initially. It is lexically assigned and says nothing about the semantic interpretation of plural. Usually it is interpreted as associative plural, as a noun as high in animacy-definiteness as 'father' tends to yield associative plural interpretation, as in SJ. By contrast, in (40b) the level high tone shifts toward the morpheme boundary between the stem and the plural suffix *-taa*. This tone shift induces the additive plural interpretation. Kibe et al. (2019) note that the tone shift is only available for address nouns, i.e. elder kin terms and proper names. This restricted distribution is reasonable, given that it is for this class of nouns that the additive plural reading is a highly marked choice. The tone shift reflects the functional markedness of the additive plural interpretation for address nouns.

4.3 Plural marking as atypical pluralization

A number of Japonic languages like the Yuwan dialect of Amami Ryukyuan (Niinaga 2014), the Shiiba dialect of Miyazaki Prefecture and the Osaka dialect (Kanbayashi 2017) allow inanimate nouns to take the same suffix as that used for pluralizing animate nouns, but the semantic effect is not that of pluralization (either associative or additive). The exact nature of these atypical plurals is still unclear, and may correspond to notions of 'approximate plural' (Corbett 2000), 'similative plural' (Daniel and Moravcsik 2013), etc. Interested readers are referred to Niinaga (2020)

for a more detailed study on these atypical plurals in Japonic languages. Below I take up only one of the most recurrent functions of atypical plurals in Japonic, the one called the exemplificative in Japanese linguistics.

In the Senshu dialect of Osaka (Oshima 2018), the plural suffix *-ra* encodes plurality, associative or additive, when attached to human nouns and animal nouns like those designating domesticated animals, as in *watai-ra* (1-PL) 'we/us', *sensei-ra* (teacher-PL) 'teachers; teacher and associates', *inu-ra* (dog-PL) 'dogs', etc. The suffix *-ra* also occurs on inanimates, but in this case it never yields associative or additive plural meaning as expected of animate nouns. Rather, the semantic effect is what Oshima calls 'exemplification', whereby the *-ra* plural denotes a list of objects which is exemplified by the referent (*kyookasyo-ra* 'textbook-PL ('a textbook and things like that').

(41) *kyookasyo-ra hayaku katazukee=ya*.
　　 textbook-PL　quickly　put.away.IMP=SFP
　　 'Put away the textbook and things.'

Here, *kyookasyo-ra* denotes a set of objects which includes learning materials like pencils, erasers, notebooks, etc., and the textbook is just an exemplar of the set.

A similar phenomenon is found in other Japonic languages like the Shiiba dialect of Miyazaki Prefecture, where the plural suffix *-domo* may be used to encode the category of exemplificative.

(42) *sumoo-domo miyotta*.
　　 Sumo-PL　　watch.IPFV.PST
　　 '(I) used to watch Sumo wrestling and so on.'

In (42) *sumoo-domo* denotes a set of TV programs of which a program on Sumo wrestling is taken up as an example. It is impossible to stipulate the other members of the denotation in an exact way.

5 Conclusions

The present chapter provides an overview of the number system across the Japonic family, focusing mainly on the number system of SJ with reference to dialectal variation. Most Japonic dialects and SJ share a number of features such as the singular-plural distinction, absence of formal contrast between associative and additive plural, etc. An important feature of Japonic is the address noun category which assimilates to pronouns in terms of obligatoriness of plural marking and its semantic interpretation.

There is also conspicuous dialectal variation with regard to, among other things, the presence of the dual category and/or the clusivity distinction (Ryukyuan), the explicit plural marking for generics as well as in relational possessives and intentional predicate constructions (Shiiba dialect of Miyazaki Prefecture) and the availability of a disambiguation strategy whereby tone shift induces the additive plural reading (Yoron dialect of Amami Ryukyuan). Also, the exemplificative plural is reported in various Japonic languages (Osaka Japanese, Shiiba dialect of Miyazaki Prefecture, etc.).

Abbreviations

ACC	accusative
CLF	classifier
COP	copula
CVB	converb
DAT	dative
DSC	discourse marker
EMP	emphatic
GEN	genitive
GER	gerundive
HS	hearsay
IMP	imperative
INT	intentional
IPFV	imperfective
LOC	locative
NLZ	nominalizer
NOM	nominative
NPST	nonpast
PL	plural
PRF	perfect
PROG	progressive
PST	past
Q	question
RSL	resultative
SEQ	sequential
SFP	sentence-final particle
THM	thematic vowel
TOP	topic

References

Bhat, D. N. S. 1994. *Pronouns*. Oxford: Oxford University Press.

Daniel, Michael & Edith Moravcsik. 2013. The associative plural. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/feature/36A#2/28.3/148.4.

Ghomeshi, Jila. 2003. Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57. 47–74.

Iljic, Robert. 1994. Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32. 91–116.

Irwin, Mark & Mattew Zisk. 2019. *Japanese linguistics*. Tokyo: Asakura Shoten.

Iwasaki, Shoichi. 2013. *Japanese (revised edition)*. Amsterdam: John Benjamins.

Kanbayashi, Aoi. 2017. Kansai hoogen ni okeru setsubiji 'ra' [The suffix 'ra' in Kansai dialects] *Handai Shakai Gengogaku Nooto* 15. 59–71.

Kazama, Shinjiro. 2003. Do the Three Groups of the "Altaic" Languages (Turkic, Mongolic, and Tungusic), as well as Korean, really Resemble Japanese Grammatical Structure?: An Attempt at a Contrastive Grammar Analysis. In Alexander Vovin, Toshiki Osada & Kerri Russel (eds.), *Nihongo keitooron no genzai* [Perspectives on the Origins of the Japanese Language], 249–340. Kyoto: Nichibunken.

Kibe, Nobuko, Hajime Oshima & Masahiro Yamada. 2019. Plural Forms in Yoron-Ryukyuan and Address Nouns in Ryukyuan Languages. *Japanese/Korean Linguistics* Vol. 25, Stanford, CSLI Publications.

Kurafuji, Takeo. 1999. Japanese Pronouns in Dynamic Semantics: The Null/Overt Contrast, PhD thesis, Rutgers University.

Kurafuji, Takeo. 2004. Plural morphemes, definiteness, and the notion of semantic parameter. *Language and Linguistics* 5. 211–242.

Kwon, Song Nim & Anne Zribi-Hertz. 2004. Number from a syntactic perspective: Why plural marking looks 'truer' in French than in Korean. In Oliver Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics* 5. 133–157.

Nakanishi, Kimiko & Satoshi Tomioka. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13. 113–140.

Niinaga, Yuto. 2014. A grammar of Yuwan, a northern Ryukyuan language. PhD thesis, University of Tokyo.

Niinaga, Yuto 2015. Kudakajima hoogen no kan'i bunpoo [A sketch grammar of the Kudakajima dialect]. *Ryuukyuu Shogo Kijutsu Bunpoo* vol. 2. 86–104.

Niinaga, Yuto. 2020. Kita Ryukyu Amami Oshima Yuwan Hogen no meishi daimeishi hukusuukei no kinoo to tsukaiwake [Non-number Values of Nominal and Pronominal Plural Forms in Yuwan in Crosslinguistic Perspective]. *Gengo Kenkyu* 157. 71–112.

Nomoto, Hiroki. 2013. Definite-like plurals. ms.

Oshima, Hajime. 2018. Osaka senshu hoogen ni okeru 'ra' no hukusuu sei [Plurality of 'ra' in the Senshu dialect of Osaka]. *Proceedings of the 157th meeting of Linguistic Association of Japan*. 136–141.

Pellard, Thomas. 2015. The linguistic archeology of the Ryukyu Islands. In Patrick Heinrich, Shinsho Miyara and Michinori Shimoji (eds.), *Handbook of the Ryukyuan Languages*, 13–37. Boston: Mouton de Gruyter.

Sasaki, Kan. 2018. Chibaken minami boosoo shi miyoshi hoogen ni okeru shijisi to daimeishi [Demonstratives and pronouns in the Miyoshi dialect of Minami Boso City of Chiba Prefecture]. Paper presented at the collaborative research project meeting of Endangered Dialects in Japan, NINJAL.

Shimoji, Michinori. 2010. The Ryukyuan languages: an introduction. In Michinori Shimoji & Thomas Pellard (eds.), *An introduction to Ryukyuan languages,* 1–11. Tokyo: ILCAA.

Tamamura, Fumio. 1985. *Goi no kenkyuu to kyooiku* [Research on vocabulary and education]. Tokyo: National Institute for Japanese Language.

Yamada, Minoru. 1985. *Yoron hoogen taigen no gohoo* [Nominal syntax in Yoron]. Tokyo: Daiichi Shobo.

Yokoyama, Akiko. 2017. Ryuukyuu Okinoerabujima hoogen no bunpoo [A Grammar of Okinoerabu-Kunigami Dialect, Ryukyuan]. PhD thesis, Hitotsubashi University.

William A. Foley

15 Number in the Languages of the Lower Sepik Family

Abstract: For the six languages of the Lower Sepik family, number is a pervasive and elaborated grammatical category, perhaps on a scale that few other language families can match. The number of number distinctions varies from language to language and word type to word type. The most elaborate distinctions are in the pronominal systems, where the languages commonly distinguish four numbers, singular, dual, paucal (from three to about seven) and plural. Such a four way number distinction for nouns is also attested in some languages of the family. These are concordial languages as well, so number marking carries a high grammatical functional load, manifested through agreement for number (and class for a subset of the languages), as verbs must agree with their core arguments and modifiers with their heads.

1 Overview

New Guinea is far and away the most linguistically complex region of the world; in an area of about 900,000 km² or about the size of the Australian state of New South Wales is found over a thousand languages, about the same total as all of the spoken indigenous languages of the Western Hemisphere, or about 20 % of all the languages currently spoken worldwide. Language diversity is also extremely high; on current knowledge and classifications there are over fifty language families and isolates spoken in the region, with a consequent very high degree of typological heterogeneity as well. This diversity is not evenly distributed in the region, being especially focused in the northern lowlands of the main island and here particularly in the region of the drainage basin of the Sepik and Ramu rivers, an area of around 100,000 km², roughly the size of the US state of Kentucky, with sixteen language families and isolates, comprised of well over a hundred languages. Nowhere else on earth is such an efflorescence of languages compressed into such a small area, though aboriginal California, the Caucasus and the Himalayan region come close. One of the larger language families in this region is the Lower Sepik-Ramu Family, which consists of a number of quite typologically varied sub-families (Foley 2005, 2017a), one of which is my concern here, the Lower Sepik Family. This consists of six languages, all of which I have done fieldwork on, subclassified as follows (current potential numbers of speakers in parentheses under the language name; these are undoubtedly seriously inflated in most cases due to the rapid and ongoing language shift in the region):
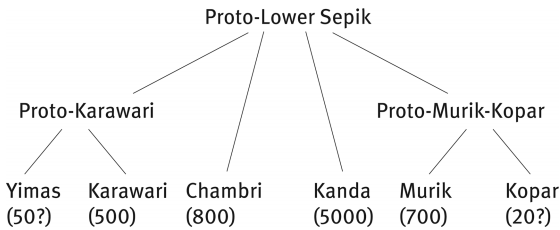
Fig. 1: The Lower Sepik Family.

Unlike many of their cousins in other sub-families of the Lower Sepik-Ramu Family (Foley 2017a), Lower Sepik languages are extremely complex morphologically. Typologically they are all head marking (Nichols 1986) and agglutinative, typically highly developed in that feature like Kanda, but in addition some like Yimas are polysynthetic, due to extensive incorporation (Foley 2017b). All six languages mark number and person of core arguments by affixes to the verb and all also inflect nouns for number. In addition, all of the languages, save Murik and Kopar, and certainly Proto-Lower Sepik, have a system of noun classes, varying from around seven in Chambri to ten in Yimas. In the languages with noun classes, class and number are expressed in portmanteau morphemes fused with the root, and all nouns are inflected for number. Number marking is obligatory and productive, but it is not regular. The allomorphy of number marking depends on noun class; dual is normally more regular than plural marking.

The set of number distinctions varies from language to language and word type to word type. The most elaborate distinctions are in the independent pronominal systems, where the languages generally distinguish four numbers, singular, dual, paucal (from three to about seven) and plural, although not necessarily in all persons (in Kanda this is found only in first person). No contemporary language productively makes such a four way distinction for most nouns, though nineteenth and early twentieth century Murik did (Schmidt 1953). Contemporary Murik only now seems to make this four way number distinction for those nouns with high animate and particularly human referents, while Chambri has it vestigially for a handful of nouns (Pagotto 1976); clearly Proto-Lower Sepik did as well. Yimas, Karawari, and Chambri inflect nouns for three numbers, singular, dual and plural, and Kanda and Kopar for only two, singular and plural (and this not even obligatorily in Kopar). There are a number of nouns with mass semantics that are invariable, e.g. 'blood' Chambri *yari,* Yimas *yat,* Murik *yaran;* 'water' Chambri, Yimas, Murik, Kopar *arəm;* 'lime powder' Chambri *ayər,* Yimas *awi,* Murik *ayr,* Kanda *awer,* and a few collective nouns that might be surprising: 'fire' Chambri *ayir,* Yimas *awt,* Murik *aur,* Kopar *awr;* 'mosquito' Chambri *naŋgun,* Yimas *naŋkun,* Kopar *naŋgit;* 'hair' Chambri *wabri,* Yimas *wapwi;* 'louse' Yimas *kuran.* The marking of number for core arguments of the verb is even more varied. All six languages are heavily head marking and typically lack case marking for core arguments, which are therefore indicated by verbal

affixes. In spite of its depauperate marking for number on nouns in comparison to its sisters, Kopar specifies four numbers for core arguments marked on the verb. In both Murik and Kopar there are distinct portmanteau affixes for this function which express both person, first, second and third, and number, singular, dual, paucal and plural (Foley 2016). Yimas also distinguishes four numbers for core arguments, but lacks a distinct series of portmanteau affixes for paucal number for first and second person. Rather it uses the corresponding plural series of affixes in combination with a generalized paucal suffix, and Kanda appears to behave similarly, but only for the person in which it distinguishes paucal number in the independent pronouns, i.e. first person. In second and third person, Kanda only distinguishes two numbers, singular versus plural. Finally, Karawari and Chambri straightforwardly express a three way number contrast for core arguments, singular, dual and plural.

Lower Sepik languages are concordial languages and make heavy use of agreement to signal grammatical relations. Constituency is very weakly developed and they are prototypical cases of non-configurational languages (Hale 1983). Modifiers of a noun need not form a constituent with it and can be floated away, but they must agree in noun class and number with it, as in this example from Kanda:

(1) *paruŋgli*   ***kle***   *ami-na-**ŋklea** **kl**-erəm   kupa-**ŋklea***
    betelnut.**III.PL** these.**III.PL** 1SG-POSS-**III.PL** **III.PL**-three big-**III.PL**
    'these three big betelnuts of mine'

And with the floated modifier in this Yimas example:

(2) *parwa-n*   *kanta kamta-k-**wa**   ta-pu-tay-kiak-rm*
    dock.**IX.SG**-OBL but   empty-IRR-**IX.SG** NEG-3-see-IRR-3DL.NOM
    kay
    canoe.VIII.SG
    'but they both didn't see a canoe at the empty dock'

These are concordial languages, and generally 'verbal number' seems to be absent. It is clearly absent in the language most studied, Yimas (Foley 1991), but it is found in at least one language, nineteenth and early twentieth century Murik; this feature needs more study.

In the following sections the examples will be drawn more from Yimas and the language for which the data are fullest, though data from the other five languages will also be presented.

2 Pronominal, nominal and verbal number

2.1 Pronominal number

As mentioned earlier, the maximal distinctions in number, singular, dual, paucal and plural are made in the independent pronouns. These are always independent and contrastive and never take case marking in these heavily head marking languages. The function of case marking to signal grammatical functions is restricted to the bound pronominals. Murik and Kopar independent pronominals have the full set of contrasts across all three persons:

Tab. 1: Murik Independent Pronouns.

	SG	DL	PC	PL
1	ma	ga-i	ag-i	e < *a + i
2	mi	ga-u	ag-u	o < *a + u
3	mən	məndəb	məŋgə	mwa

Yimas is similar though it neutralizes the contrast between first and second person in the paucal number (Yimas has a great number of underlyingly vowel-less syllables and even words, as the epenthetic /ə/ is completely phonologically predictable; the orthography reflects this (Foley 1991); the status of this vowel in the other languages is not as straightforward and hence it is here represented orthographically):

Tab. 2: Yimas Independent Pronouns.

	SG	DL	PC	PL
1	ama	kapa	paŋkt	ipa
2	mi	kapwa		ipwa
3	mn	mrm	mŋkt	mum

Karawari only distinguishes three numbers:

Tab. 3: Karawari Independent Pronouns.

	SG	DL	PL
1	ama	kapa	apia
2	mi	kupa	ipa
3	mən	ŋkri	mpu

Chambri and Kanda are particularly interesting; both distinguish all four numbers in first person, but Chambri only three numbers and Kanda only two in second and third:

Tab. 4: Chambri Independent Pronouns.

	SG	DL	PC	PL
1	*ami*	*kəpi*	*ipəŋk*	*ipi*
2	*mi*	*kəbwi*		*ibwi*
3	*mən*	*mənəmp*		*mump*

Tab. 5: Kanda Independent Pronouns.

	SG	DL	PC	PL
1	*ame*	*paŋke*	*paŋkər*	*ipe*
2	*mi*			*ipwe*
3	*mən*			*pun*

Generally, the morphological formation for number of these independent pronouns is unique and does not overlap with that of nouns, with the exception of third person. At least in the case of Yimas, the so called third person pronouns are actually deictics, the set illustrated in Table 2 being the near distal series built on a stem in *m-*, to which are added suffixes belonging to the nominal set (except for paucal, which is unique, as nouns and nominal modifiers are not inflected for a distinct paucal number in Yimas. For example, the suffix *-rm* for the third dual form is also employed on certain classes of nouns to indicate dual number, 'star' SG *awak* DL *awakrm,* and *-um* for third plural to indicate plural, 'woman' SG *narmaŋ* PL *ŋaykum*. A similar pattern is found in Chambri: 'husband' SG *manan* DL *mananəmp* PL *mananamp*. Within the independent pronoun systems a number of formatives can be identified across the six languages: a paucal formative *-ŋk/ŋg,* a dual stem *ka-/ga-,* a high front vowel for first non-singular and a corresponding back vowel or semi-vowel for second person non-singular.

When we come to the bound verbal personal pronominals, only Murik and Kopar again have forms which contrast all four numbers in all three persons. Furthermore, the bound verbal pronominals in these languages exhibit a split ergative-accusative system, with the local persons, first and second, operating on a nominative-accusative basis or three-way, i.e. contrasting subjects of intransitive verbs with subjects of transitive verbs and with objects of transitive verbs, as Murik does for the singular number in Table 6 below (in bold). The non-local person, the third person, works on an ergative-nominative basis.

Note that the three way contrast is only in the singular. The non-singular numbers collapse nominative and ergative; in other words, there is strictly a nominative-accusative distinction. The non-local person is straightforwardly ergative-nominative (Table 7).

The INV allomorphs for the ergative in the non-singular refer to whether they are being used in a verb with inverse inflection, a feature of all six languages which adds greatly to the complexity of their verbal morphology, which I will discuss below (see also Foley (2016) for details).

Tab. 6: Murik Local Bound Pronominals.

		NOM	ERG	ACC
1	SG	*ma-*	*a-*	*aŋa-*
	DL	*age-*		*ŋe-*
	PC	*agi-*		*ŋi-*
	PL	*e-*		*ŋe-*
2	SG	*me-*	*Ø*	*ana-*
	DL	*ago-*		*ŋo-*
	PC	*agu-*		*ŋu-*
	PL	*o-*		*ŋo-*

Tab. 7: Murik Non-local Bound Pronominals.

		NOM	ERG
3	SG	*o-* / *Ø-*	*Ø-*
	DL	*bo-*	*bo-* / *mb-* (INV)
	PC	*d-*	*bo-* / *ŋg-* (INV)
	PL	*g-*	*bo-* / *mbu-* (INV)

Yimas is less elaborate here than Murik and Kopar. It too has a split ergative-accusative system along the lines of local versus non-local persons (although the local persons are three way distinctive, with the exception of first dual), but distinguishes only three numbers in its bound pronominals for local persons.

Tab. 8: Yimas Local Bound Pronominals.

		NOM	ERG	ACC
1	SG	*ama-*	*ka-*	*ŋa-*
	DL	*kapa-*	*ŋkra-*	*ŋkra-*
	PL	*ipa-*	*kay-*	*kra-*
2	SG	*ma-*	*n-*	*nan-*
	DL	*kapwa-*	*ŋkran-*	*ŋkul-*
	PL	*ipwa-*	*nan-*	*kul-*

Whereas, surprisingly, and in contrast to Chambri and Kanda in Tables 4 and 5 above, in the non-local person, all four numbers are contrasted (Table 9).

Tab. 9: Yimas Non-local Bound Pronominals.

		NOM	ERG
	SG	*na-*	*n-*
3	DL	*impa-*	*mp-*
	PC	*kra-*	*ŋkl-*
	PL	*pu-*	*mpu-*

Yimas generally lacks distinct bound paucal pronominals for the local persons; the only exception is the nominative form, used for the subjects of intransitive verbs, for which there is a distinct paucal prefix, but as with the independent pronouns, it is ambiguous as to first or second person referents:

(3) *paŋkra-wa-t*
     1/2PC.NOM-go-PERF
     'we/you few went'

In the case of ergative or accusative pronominals, no such form exists; one must use the corresponding plural bound local pronominals in combination with an invariable paucal suffix *-ŋkt,* essentially 1/2PC:

(4) a. *pu-kay-cay-c-ŋkt*
        3PL.NOM-1PL.ERG-see-PERF-PC
        'we few saw them' (not 'we saw them few')

     b. *pu-kra-tay-c-ŋkt*
        3PL.NOM-1PL.ACC-see-PERF-PC
        'they saw us few' (not 'they few saw us')

Chambri and Karawari overall seem to pattern much like Yimas, but on current data available it is not known whether they exhibit paucal marking for bound pronominals. Kanda does, and again not surprisingly, distinguishing all four numbers only in first person, following the pattern of the independent pronouns, but contrasting only singular versus plural in non-first persons. For first person, with the intransitive verb 'bathe', we find:

(5) a. *mən-ta-tupwala-ka-nəm*
        1-AUX-bathe-PRES-1SG
        'I am bathing',

     b. *mən-ta-tupwala-ka-ŋke*
        1-AUX-bathe-PRES-1DL
        'we both are bathing',

c. *mən-ta-tupwala-ka-ŋkər*
1-AUX-bathe-PRES-1PC
'we few are bathing',

d. *mən-ta-tupwala-ka-nte*
1-AUX-bathe-PRES-1PL
'we many are bathing'

But for third person:

(6) a. *an-ta-tupwala-ka-n*
3-AUX-bathe-PRES-3SG.NOM
'he is bathing'

b. *an-ta-tupwala-ka-num*
3-AUX-bathe-PRES-3PL.NOM
'they (two/few/many) are bathing'

And with bound pronominals of transitive verbs in some combinations, all number distinctions in third person are erased:

(7) a. *mən-ta-ti-ka-ŋkər*
1-AUX-hit-PRES-1PC
'we few are hitting him/them'

b. *pwan-ta-ti-ka-ŋkər*
3.ERG-AUX-hit-PRES-1PC
'he/they are hitting us few'

Such neutralization never occurs with first person.

Both the demonstrative pronouns (i.e. deictics) and the interrogative pronoun for 'who' in Yimas and Murik (Schmidt 1953), like the personal pronouns, distinguish all four numbers. Here are the forms for the proximal demonstrative 'this' and 'who' in both languages:

Tab. 10: Demonstrative and Interrogative Pronouns in Yimas and Murik.

	Yimas		Murik	
	PROX	**'who'**	**PROX**	**'who'**
SG	*na-k*	*naw-n*	*e-wa*	*mən-amena*
DL	*impa-k*	*naw-rm*	*e-bwa*	*məndəb-amena*
PC	*kra-k*	*naw-nkt*	*e-ra*	*məŋgəŋ-amena*
PL	*pu-k*	*naw-m*	*e-ga*	*mənduŋ-amena*

2.2 Nominal number

Number inflection is an obligatory feature of all nouns except mass nouns in all Lower Sepik languages but Kopar and Murik, in which it is optional, but highly preferred for nouns with higher animate or human referents. In the four languages preserving noun classes, Yimas, Karawari, Chambri and Kanda, number is indicated in a portmanteau fashion with noun class, in a manner very similar to how these two categories are indicated in Bantu languages. However, unlike Bantu languages in which each noun class selects a unique marker for both numbers, singular and plural, in Lower Sepik languages there is some neutralization in the non-singular numbers, but the degree of this varies widely from language to language, more extensive in Chambri and Karawari for instance than Yimas. Still, the class contrast remains robust. In Yimas, there is a myriad of allomorphs for number, some of them found across more than one class. For instance, the suffix *-i* marks the plural of some nouns across several classes. However, regardless of the shared allomorph for the marking of nominal number, the underlying class distinctions are still salient, apparent, for example, in the corresponding concordial suffixes for adjectives: class I *-ump* versus class II *-nput* versus class VII *-ra*. Hence *-i* cannot be analyzed as a maker of plural number; it is a fused class-number marker, with its distribution selected by class, albeit overlapping across several classes.

Commonly, the number distinctions for nouns correspond to that in the third person or demonstrative pronouns, but there are two exceptions. Kopar distinguishes four numbers in its third person pronouns, yet only optionally distinguishes singular and plural for nouns, while Yimas distinguishes four numbers in pronouns and demonstratives, but only three, singular, dual and plural, for nouns.

Let me first consider the simpler case of Kopar. In Kopar an uninflected noun is vague for singular or plural number, but plural can be optionally marked overtly by a particle, *okumbi*

(8)   SG   PL
     'tooth'   *asirap*   *asirap okumbi*
     'tree'   *iror*   *iror okumbi*

However, there are some irregular plural allomorphs for a few nouns referring to humans, particularly kin:

(9)   SG   PL
     'person'   *nor*   *nor-əmbak*
     'older brother'   *kakan*   *kakar-əmbak*
     'younger brother'   *ram*   *ram-taŋgar* (or *ramar-əmbak*)
     'older sister'   *maman*   *mama-taŋgar* (or *mamar-əmbak*)

In Kopar, plural marking, although grammatical, is rarely if ever used for nouns with non-human referents, but for nouns with human referents it is common, even preferred.

Murik preserves what unquestionably had been the situation in Proto-Lower Sepik, full inflection of nouns for all four number contrasts. Examples are (Schmidt 1953):

(10)

	SG	DL	PC	PL
'person'	*nor*	*norəmbo*	*normoara/norəgə*	*normot*
'house'	*iran*	*irambo*	*iramoara*	*iranmot*
'dog'	*pitak*	*pitakəmbo*	*pitakmoara*	*pitakmot*
'tree'	*yarar*	*yararəmbo*	*yararmoara*	*yararmot*

The *-(ə)gə* allomorph for PC in 'person' is of Proto-Lower Sepik ancestry and seems restricted to nouns with human referents, hence 'child' *gran* PC: *grangə*. None of the other four noun classifying languages make such a four way distinction for number, but paucal number is found vestigially for a few nouns in Chambri (Pagotto 1976), and one of the allomorphs for paucal particularly for nouns with human referents is *-aŋk*, cognate with Murik *−(ə)gə*:

(11)

	SG	PL	PC
'man'	*noranan*	*noramp*	*norakraŋk*
'daughter'	*pisanim*	*pisamtamp*	*pisamtaŋk*

Otherwise paucal number in Chambri is marked by a suffix *-reb* usually added to the plural forms of nouns:

(12)

	SG	PL	PC
'mountain'	*bawi*	*bawikas*	*bawikas-reb*
'tree'	*yuwan*	*yuwari*	*yuwa-reb*
'taro'	*warimai*	*warimaikas*	*warimaikas-reb*
'fish'	*kawi*	*kawikas*	*kawikas-reb*

There are also nouns of human or animal referents with paucal forms in Karawari as well, but in this language these are found where the dual has been lost, so the paucal ranges from two to around seven: 'female cassowary' SG: *awanma* PC: *awanmənti* PL: *awaws* (Telban 1992).

Number marking for nouns in the four noun classifying languages is much more formally elaborated than in Murik or Kopar, and I will illustrate here mostly with data from Yimas, a typical and largely conservative language, with mentions of innovations in the other three languages. Nominal classification along the lines of Yimas was unquestionably a feature of Proto-Lower Sepik but has been lost in Mu-

rik-Kopar, though relics of it remain to point to its earlier existence; for example, the numerals for 'two' and 'three', Murik *kobo* 'two', *keroŋgo* 'three', Kopar *kombari* 'two' *keremən* 'three', all exhibit a fossilized prefix *k-* the verbal agreement marker used on these two numerals for class VI in Yimas and the corresponding classes in the other three languages. Number marking in Yimas is very complex and often suppletive or irregular, but is determined by noun class membership. The dual inflection is more regular and exhibits less allomorphy than the plural, and a good deal of the intricacy in plural inflection is due to the collapse and loss of the old paucal number, so that the synchronic allomorphs for plural in various noun classes can reflect older paucal or plural markers. A somewhat similar development has occurred in Kanda, in which the dual as well as the paucal number has been lost, so that plurals in that language can reflect older plurals, paucals or duals. Yimas possesses ten distinct noun classes, one of which is divided into two sub-classes. In addition, there are a few common nouns like *nam* 'house', *num* 'village' and *arm* 'water' that form separate classes all by themselves. The principles that determine noun classification are twofold: semantic and phonological. There are four classes determined semantically: nouns referring to male humans, female humans, animals and culturally useful and important plants. The other six classes are determined phonologically by the final underlying one or two segments of the root: roots ending in *ŋk,* roots ending in *mp* (for each of these two classes the final stop deletes in word final position, that is, in the singular, as marking of singular is zero), roots ending in *i* or *y,* roots ending in *w* (the *w* again deletes in word final position, i.e. in the singular), roots ending in *uk* or *uŋk*, and finally a general class containing roots ending in any other consonant or consonant cluster as well as some roots ending in *i* or *y* not belonging to the previous small class mentioned.

The number inflection of a noun is determined by the noun class to which it belongs, but the allomorphy can be quite complex and irregular, especially for the semantically determined classes and the general phonological class; it is not the case, particularly for plural inflection, that there is only one allomorph, for often there are several. Consider class I, the one with male human referents. There is both suppletion for plural here ('man' SG: *panmal* PL: *payum*, 'boy' SG: *kaylakn* PL: *kumpwi*) as well as back formation, where the singular is derived from the plural by the unusual step of adding a masculine singulative nominalization suffix, 'person' PL: *namat* SG: *namarawt,* 'adolescent boy' PL: *taŋkat* SG: *taŋkarawt* (final *t* corresponds to *r* intervocalically), but even with that there are no less than four allomorphs of the plural suffix lexically and partially phonologically determined, *-um, -am, -i* and *−(n)tt ~ -(n)ct* (the later arising from palatalization):

(13)

	SG	PL
'father'	*apwi*	*apwiam*
'brother'	*matn*	*matnum*
'father-in-law'	*macawk*	*macawkwi*
'old man'	*apanwakn*	*apanwakntt*

The formation of dual is much simpler: except for the single example of partial suppletion ('boy' SG: *kaylakn* DL: *kaymampan), the* dual is always marked by the suffix *-rm,* with an intrusive *t* (palatalized to *c* after a high front vowel or semivowel) between the suffix and the final segment of the root if it is anything other than *k*:

(14)   SG   DL
    'father-in-law'   *macawk*   *macawkrm*
    'father's sister's husband'   *apuk*   *apukrm*
    'brother'   *matn*   *matntrm*
    'mother's brother'   *away*   *awacrm* (palatalization)

Class II, nouns with female human referents, is similar, with both suppletion ('woman' *narmaŋ* PL: *ŋaykum)* and a number of allomorphs for plural:

(15)   SG   PL
    'sister'   *apak*   *apaki*
    'mother'   *ŋayuk*   *ŋaykumpam*
    'father's sister'   *ŋaki*   *ŋakŋct* (palatalization)

This class is unusual in having a distinct female marker *-maŋ,* often deriving nouns in this class from class I nouns by its addition. When the plural is formed this suffix is dropped and the plural marker is invariably *-(m)put:*

(16)   SG   PL
    'mother-in-law'   *macawkmaŋ*   *macawkmput*
    'adolescent girl'   *taŋkarmaŋ*   *taŋkarmput*
    'daughter-in-law'   *pranmaŋ*   *pranput*

But dual inflection is surprising for this class compared to class I, as there are two allomorphs: forms ending in *-maŋ* in the singular take *–(m)prum* after dropping the *-maŋ,* but those without it take the same dual suffix as class I nouns:

(17)   SG   DL
    'sister'   *apak*   *apakrm*
    'father's sister'   *ŋaki*   *ŋakcrm*
    'mother-in-law'   *macawkmaŋ*   *macawkmprum*
    'adolescent girl'   *taŋkarmaŋ*   *taŋkarmprum*

Class III nouns with referents of animals are simpler and mostly behave like class I, but with less complexity. The plural marker for most nouns of this class is *–i,* sometimes with an intrusive *w* or the dropping of final *n* of the root:

(18)　　　　　　　　SG　　　　PL
　　　'cassowary'　　*awa*　　　*awawi*
　　　'crocodile'　　*manpa*　　*manpawi*
　　　'dog'　　　　　*yura*　　　*yuray*
　　　'pig'　　　　　*numpran*　*numpray*

However, there are a few nouns that take a plural ending typical of the phonologically general class V, *-ŋkat,* which by comparative evidence with Murik and Chambri almost certainly goes back to an old paucal suffix:

(19)　　　　　　　　SG　　　　　PL
　　　'fish species'　*wantat*　　*wantarŋkat*
　　　'wallaby'　　　*warkapwi*　*warkapwiŋkat*

The dual of class III is very straightforward, as all nouns take the same dual suffix *-ntrm* (note that this form differs from classes I and II in the homorganic nasal and the obligatory *t):*

(20)　　　　　　　　　　SG　　　　DL
　　　'cassowary'　　　*awa*　　　*awantrm*
　　　'dog'　　　　　　*yura*　　　*yurantrm*
　　　'bird of paradise'　*namakat*　*namakatntrm*
　　　'wallaby'　　　　*warkapwi*　*warkapwiɲcrm* (palatalization)

The nouns of class IV, the last class of semantically determined classes, with the names of culturally important plants, is quite different from the previous three (which probably all go back to a single noun class in Proto-Lower Sepik that split into three in Yimas) and in many ways shares features with the phonologically determined classes to be considered next. The roots of nouns in this class normally end in *-um,* so in a sense class membership is both semantically and phonologically determined. With just two exceptions which take the likely older paucal marker *-ŋkat,* nouns in this class take *-ŋi* for plural after dropping the final *m* of the root:

(21)　　　　　　　　　　SG　　　　PL
　　　'type of sago palm'　*tmarum*　*tmaruŋi*
　　　'coconut palm'　　　*irp(u)m*　*irpuŋi*
　　　'type of tree'　　　*nŋkrum*　*nŋkruŋi*
　　　'sago palm'　　　　*tnum*　　*tpwi* (suppletive)

All nouns in this class take *-ul* for dual, where the *l* is a palatalized lateral, the reflex of an older dual marker **-ri* still preserved in Karawari, Kanda and Chambri:

(22)   SG   DL
    'coconut palm'   *irp(u)m*   *irpmul*
    'sago palm'   *tnum*   *tnumul*
    'type of tree'   *nŋkrum*   *nŋkrumul*

The marker *-l* is the typical maker for dual in all the phonologically determined classes except the general one, class V.

Noun class V is by far the largest in the language and contains all nouns that neither fall into the semantic categories of classes I–IV nor end in the final segments of classes VI–X. As might be expected with such a large and heterogeneous class the marking for plural is quite diverse, some nouns taking older paucal markers, *-(ŋ)kat* or *-ra* (compare Murik), and others plural markers, *-i* or *-t*, though *-(ŋ)kat* is by far the most common plural suffix in this class. Which is chosen is lexicalized and cannot be determined phonologically:

(23)   SG   PL
    'arm band'   *nawran*   *nawray*
    'liver'   *tampan*   *tampanra*
    'fish species'   *tamun*   *tamut*
    'owl'   *wamun*   *wamura*
    'fish species'   *pian*   *piankat*

Roots in class V have these four distinct plural allomorphs, and their possibilities can be classified according to the final segments of the roots they occur with (see Table 11).

Tab. 11: Allomorphy for Plural Marking for Class V Nouns in Yimas.

PL suffix	final segment of root
-ŋkat ~ -kat (after nasals) ~ *-at* (after *k*)	*p, k, m, n, ŋ, r, l, i, y, c*
-i ~ -y	*k, n*
-ra	*n*
-t	*n, nt, r*

Dual marking for this class is much more straightforward; it is always a variant of *-rm*. After *k* and the nasals it is just *-rm,* though the *r* fortifies to *t* after a nasal, but elsewhere it is in the longer form *-ntrm* found in class III nouns (this can be used optionally with nouns ending in *k* as well):

(24)		SG	DL
	'star'	*awak*	*awakrm* or *awakntrm*
	'tree'	*yan*	*yantm*
	'water python'	*aŋkaŋkam*	*aŋkaŋkamtm*
	'net bag'	*yakut*	*yakutntrm*
	'rope'	*awkp*	*awkpntrm*
	'nose'	*tkay*	*tkaɲcrm* (palatalization)

The other five noun classes all share with class IV the invariable dual marker *-l* from **-ri,* though they differ in the plural allomorphs selected. Class VI is far and away the most regular in Yimas (and in the other three languages which have preserved noun classification; it is exemplified in Kanda in example (1) above, where it is labeled as class III, as Kanda has fewer noun classes than Yimas), and so it is not surprising that when the noun class system was lost in Murik-Kopar, it was the forms of the numerals 'two' and 'three' that were targeted by agreement controllers from this class that survived as the invariable numerals in those two languages. The roots of nouns in this class all end in *ŋk,* with loss of the stop in final position; because singular number has no overt inflection, nouns in singular number just end in *ŋ.* In this class all nouns take the suffix *-i* for plural, unquestionably an old plural affix, and *-l* for dual:

(25)		SG	DL	PL
	'head'	*yampaŋ*	*yampaŋkl*	*yampaŋki*
	'green tree frog'	*akrŋ*	*akrŋkl*	*akrŋki*
	'glans penis'	*karŋ*	*karŋkl*	*karŋki*
	'coconut'	*uraŋ*	*uraŋkl*	*uraŋki*

As mentioned earlier, Kanda has lost the dual number entirely, and for this class the dual suffix is now the plural marker in these Kanda examples:

(26)		SG	PL
	'tooth'	*sisiŋ*	*sisingli*
	'betel nut'	*paruŋ*	*paruŋgli*
	'coconut'	*waŋ*	*waŋgli*

Class VII parallels class VI, but is not quite as regular. Roots of nouns in this class also terminate in a homorganic nasal plus stop cluster, which again simplifies to just the nasal in final position, but this time the cluster is *mp.* Again, the dual marker is invariably *-l,* but the plural is more complicated. The large majority of nouns take a suffix *-at,* but there are a few that like class VI occur with *-i* (often with loss of *mp* or other root simplification) and one with *-a*:

(27)

	SG	DL	PL
'basket'	*impram*	*imprampl*	*imprampat*
'arrow'	*piam*	*piampl*	*piampat*
'toe'	*nukraym*	*nuŋkraympl*	*nuŋkray*
'roofing'	*antkm*	*antkmpl*	*antiki*
'plate'	*aprm*	*aprmpl*	*apra*

Class VIII is a very small class of a little over a dozen nouns ending in a high front vowel or semivowel (other nouns ending in these segments belong to the general class V). Again the dual is invariably *-l,* and the plural is usually *-(ɲ)cmpt,* from underlying *-(n)tmpt* by palatalization after the high front segment:

(28)

	SG	DL	PL
'canoe'	*kay*	*kayl*	*kaycmpt*
'breast'	*nŋay*	*nŋayl*	*nŋayɲcmpt*
'clay pot'	*yaŋi*	*yaŋil*	*yaŋɲcmpt*

There are, however, two roots in this class that take the rare plural suffix *-a* that we saw with 'plate' above in class VII: 'flying fox' *kumpwi* PL: *kumpwia,* 'cross beam of house' *waŋki* PL: *waŋkia.* Class IX is the corresponding class of roots which end in a high back semivowel, but this can only appear as part of the diphthong *aw,* so underlyingly these roots all end in *aw.* However, both the high back vowel and the semivowel are prohibited phonotactically from occurring word finally, so that the forms of these roots in singular, which again have no overt affixation, end in *a.* For some roots the high back semivowel disappears entirely, yet for others it metathesizes to form a cluster with the preceding consonant. Again for this class the dual is invariably *-l,* and the plural is *-ut,* which is added to the stub of the root after the diphthong *aw* has been deleted:

(29)

	SG	DL	PL
'penis'	*ŋarwa*	*ŋarwawl*	*ŋarut*
'sago beetle'	*krpa*	*krpawl*	*krput*
'bark'	*yanara*	*yanarawl*	*yanarut*

Finally, class X is again quite small and developed from a split from class VI; its roots end in the high back vowel plus a velar stop or homorganic velar nasal plus stop cluster. But small though it may be, it is divided into two subclasses, one ending in *uk* and the other in *uŋk,* and this distinction shows up in the allomorphy for plural as well as in adjectival and verbal concordial affixes. For both subclasses the dual is again invariably *-l* (though with vowel harmony to insert a high back vowel between the root and the suffix), and for roots ending in *uk* the plural marker is *-at,*

while for those in *uŋk* it is *-i* (again with vowel harmony introducing a high back semivowel between the root and the suffix):

(30)

	SG	DL	PL
'bandicoot'	*awruk*	*awrukul*	*awrukwat*
'ginger'	*mawruk*	*mawrukul*	*mawrukwat*
'egg'	*awŋ*	*awŋkul*	*awŋkwi*
'headdress'	*tmpuŋ*	*tmpuŋkul*	*tmpuŋkwi*

There are also a few words in this class with irregular or suppletive plurals:

(31)

	SG	PL
'sago pancake'	*tpuk*	*tpwi*
'meat'	*naŋkpuk*	*naŋkpt*
'brain'	*kawŋŋcmpruk*	*kawŋŋcmprukwat ~ kawŋŋcmpt*

Table 12 summarizes the number inflection for the ten noun classes in Yimas and presents comparative data from corresponding classes in the other three languages (additional Chambri data from Pagotto (1976)).

In almost all cases, the singular is marked by zero, and for phonologically determined classes I have indicated the segments involved in the singular rows. Chambri and Kanda only have a single class for humans (class I), i.e. lack the gender distinction that Yimas and Karawari possess, and there does not seem to be a separate animal class either, it being divided between a general higher animate class I which includes humans and the phonologically general one. As the Yimas-Karawari sub-group is the only one that exhibits the contrast for gender, it seems highly likely that this is an innovation in this sub-group and that Chambri and Kanda here preserve the Proto-Lower Sepik system more faithfully. Current data indicate that Kanda lacks a separate noun class for plants, so this has been omitted for this language in Table 12. Overall, though, Kanda appears to be the most divergent, with new classes innovated which do not match up to those in the other three languages and cannot be treated here. I have only included classes which seem to match those in the other three languages. Kanda is a very intricate language which calls out for further research. Karawari has lost the dual number in some classes and preserved it in others (mainly only preserved it with the transparent dual marker *-ri*), while Kanda has lost it entirely. From this table we can see that the systems of Yimas and Karawari are most closely related (they form a sub-group within the Lower Sepik family) and that there are some phonological changes from Karawari to Yimas, i.e. *ri > l* and *r, s > t* in word final position.

As can be gleaned from the array of Yimas noun classes and also those in the other Lower Sepik languages, animacy plays an important part in assigning nouns

Tab. 12: Number Inflection by Noun Class in Lower Sepik Languages.

Yimas

Classes	I	II	III	IV	V	VI	VII	VIII	IX	X
SG		-maŋ		um	p, k, m, n, ŋ, nt, r, i, y	ŋk	mp	i, y	aw	uk, uŋk
DL	-rm	-rm, -prum	-ntrm	-l	-ntrm	-l	-l	-l	-l	-l
PL	-um, -am, -i, -t	-put, -i	-i, -ŋkat	-ɲi	-ŋkat, -ra, -i, -t	-i	-at, -i	-ɲcmpt	-ut	-at, -i

Karawari

Classes	I	II	III	IV	V	VI	VII	VIII	IX	X
SG		-ma		um	p, k, m, n, ŋ, nt, s, r, i, y	ŋk	mp	i, y	o, a	uk, uŋk
DL	-ŋkri	-ŋkri				-ri	-ri	-r	-i	-i
PL	-m, -ŋkar, -nti	-m, -ŋa, -ŋkar, -nti	-ya, -as	-is	-ar, -i, -ŋkar, -ya	-i	-as	-r	-i	-as, -i

Chambri

Classes	I	II	III	IV	V	VI	VII
SG		um	p, k, m, n, nt, r, i, e, y	ŋk	mp	i, e, y	u, o, w
DL	-anemp	-muri	-sim	-kri	-ari	-riri	-ri
PL	-amp, -ar, -as	-i, -e	-ar, -ri, -kas	-r	-as, -ar	-r	-r, -ar

Kanda

Classes	I	II	III	IV	V	VI
SG		*p, t, k, m, n, r, l, s, a, i*	*ŋk*	*mp*	*-l, -e*	*-we*
PL	*-m, -oli, -ŋka*	*-(l)r, -ntər, -ŋka, -oli, -e*	*-li*	*-ar, -(w)a*	*-r, -or*	*-ur*

to them. Chambri and Kanda have a distinct class for nouns that have higher animate referents, including humans, while Yimas and Karawari have split that into three: one each for male humans, female humans and animals. Further in all four languages there is one plural allomorph restricted to nouns with human referents; this is Yimas, Karawari, Kanda -*m*, Chambri -*mp* going back to Proto-Lower Sepik **-mb*. The reading of a pluralized noun depends in Yimas on whether the verb governing it is marked for paucal number or not. If it is, the noun will be interpreted as referring to a set of objects, from three to seven or so – the upper bound not being rigid, though the bottom is, as paucal directly contrasts with dual, which is strictly 'two'. This is quite different from the situation in Karawari, in which a paucal noun like *awanmənti* 'a few cassowaries' actually includes 'two' and means from 'two to about seven', a distinct dual form lacking for such nouns (Telban 1992). The interpretation of plural in Yimas is normally a larger group, over seven, but again this is not rigid; the contrast between paucal and plural is to highlight a smaller group versus a larger one, and so depending on context the boundary between them can be shifted up or down.

There are a number of nouns that resist enumeration in Lower Sepik languages and are invariable. These correspond to what may be taken as collective nouns and mass nouns. Interestingly, these forms may be morphologically singular or plural, and their corresponding agreement patterns are in accord with whatever number they are morphologically. In other words, agreement is strictly formally determined, not semantically. Consider collective nouns first. These formally are always singular and their referents typically are insects that occur in groups. Examples in Yimas include *irpulik* 'ants', *naŋkun* 'mosquitoes' and *kuran* 'lice'. One can say *naŋkun mpan* (lit. mosquito one) 'one mosquito' (although in the fens of Yimas country they never come in ones!!), but also *naŋkun muntak* (lit. mosquito many) 'many mosquitoes', but in neither case the noun can be inflected for number, albeit singular in form and all agreement is in the singular. Mass nouns In Yimas have the typical referents of homogeneous substances, but there are some which pattern as mass nouns that might be surprising, like *napnapi* 'earthquake' and *wal* 'wind', both formally singular, and *wŋcmpt* 'name', formally plural. Mass nouns which are formally singular include: *kalk* 'sago porridge', *mlm* 'feces', *icn* 'salt' *wurk* 'dirt', *anti* 'ground, land' and *ikn* 'smoke'. Formally plural mass nouns are quite common; they include: *yat* 'blood', *ɲct* 'urine', *wapwi* 'hair', *awt* 'fire', *yaki* 'tobacco', *amtra* 'food', *maramara* 'cargo, things, stuff' and *awi* 'lime powder'. Further, considering consumables, *tpwi* 'sago' is a mass noun, though formally plural, but *patn* 'betelnut' is a count noun: *patn mpa-n* betelnut.V.SG one-V.SG 'one betelnut'.

Units of time for measuring days also behave like mass nouns and cannot be directly enumerated: *ŋarŋ* 'one day removed from today, i.e. yesterday or tomorrow', *urakrŋ* 'two days removed from today', *tnwantŋ* 'three days removed from today'. Only *ŋarŋ* 'one day removed from today, i.e. yesterday or tomorrow' among these can be pluralized in a specialized sense of successive days, but in order to do so, a special pluralizing support word is necessary: *ma ŋarŋ mpŋkat-n* (lit. other

one.day.removed PL-OBL) 'on a number of successively following days'. Other types of nouns which never inflect for number are place names like *Tamprakmal* 'Yimas village', *Kapakmal* 'Wambramas village' and *Yampun* 'Sepik River' and people's personal names such as *Yakayapan, Awikamay* and *Andikapan*. Yimas possesses few abstract nouns, but those it has also never inflect for number: *upwi* 'heat', *tark* 'cold', *wapun* 'happiness', *yapan* 'sadness', *tamana* 'sickness' and *irmut* 'shame'.

Finally, as in English, nouns with generic reference can be expressed in the singular or in the plural, cf.:

(32) *patn   na-kapuc-wat*
     betelnut.SG 3SG.NOM-chew-HAB
     'He usually chews betelnut.'

(33) *tpwi   imp(a)-awkura-wat*
     sago.X.PL 3DL.NOM-gather-HAB
     'They both always gather sago.'

The former is however far more common.

Native words corresponding to group nouns like English *police* or *flock* appear to be lacking in Yimas, nor can words like *num* 'village' or *nam* 'house' be used metonymically to refer to the inhabitants. Both inflect fully for number, but irregularly, and each constitutes a noun class on its own: 'village' SG: *num* DL: *numul* PL: *nmkat*, 'house' *SG: nam* DL: *naml* PL: *nampt*. In order to express the inhabitants of a place, the place noun is suffixed by the oblique case suffix *-n* to which are added the agentive nominalization suffixes, all of which mark gender and number:

(34) a. *num-un-mat*
       village-OBL-M.PL
       'villagers' (a mixed sex group is always realized as masculine)

    b. *Chimbu-n-mat*
       Chimbu-OBL-M.PL
       'Chimbu people',

    c. *Kavieng-n-awt*
       Kavieng-OBL-M.SG
       'man from Kavieng'

    d. *Wamut-ɲum-un-(m)prum*
       section_of_river_near_Wambramas_village-village-OBL-F.DL
       'two women from Wambramas village'

    e. *Ostralya-n-maŋ*
       Australia-OBL-F.SG
       'Australian woman'

In the case of the home village the oblique suffix is unnecessary: *Tamprakmal-mat* (lit. Yimas_village-M.PL) 'Yimas people', and apparently for neighboring villages as well: *Wamur-mat* (lit. section_of_river_near_Wambramas_village-M.PL) 'Wambramas people'.

Yimas has a very rich system of nominalizations, both non-finite and finite. Action nominalizations are always non-finite and are always marked by a non-finite suffix *-ru* (with a number of allomorphs) that replaces the tense suffix of a finite verb. To the verb marked by *-ru* are added the nominalizing affixes and these are never inflected for number, occurring in one invariable form. There are four different nominalizing suffixes depending on the semantics of the action nominal:

(35) a. nominalization of words or thought (thinking is internal speech)
          *patn   wayk-r-**mpwi** pia-ka-i-c-mpun*
          betelnut.V.SG buy-NFN-talk talk.NOM-1SG.ERG-tell-PERF-3PL.DAT
          'I told them to buy betelnut'

      b. nominalization of activity
          *nam   wark-t-**nti**   tia-ka-ira-karŋkra-t*
          house enclose-NFN-act act.NOM-1SG.ERG-APPL-tired-PERF
          'I'm tired of building houses'

      c. nominalization of customary or habitual action
          *yaki   am-t-**wal**   ntak-na-k*
          tobacco.V.PL. eat-NFN-custom.V.SG leave-IMP-V.SG.NOM
          'stop smoking!'

      d. nominalization of desire
          *tpuk   am-t-**wampuŋ**   kpa-n   ama-na-t-n*
          sago.X.SG eat-NFN-heart.V.SG big-V.SG 1SG.NOM-PRES-feel-PRES
          'I really want to eat sago' ('I feel a big desire to eat sago')

Of these four nominalizers, only the last is an independent noun and when used as such can be inflected for number: 'heart' SG: *wampuŋ* DL: *wampuŋtm* PL: *wampuŋkat,* but when it is used as a nominalizer this does not appear to be possible.

Yimas has very few true adjectives, only three, *kpa* 'big', *yua* 'good' and *ma* 'other'. Many words corresponding to adjectives in English are verbs:

(36) a. *urkpwica-k-n*
          darken-STATIVE-I.SG
          'be dark (of a man)'

      b. *kalc-k-ŋki*
          harden-STATIVE-VI.PL
          'be hard (of some things)'

c. *tŋkŋt-k-nmaŋ*
become.heavy-STATIVE-II.SG
'be fat (of a woman)'

The form of these adjectives and nominalized verbs when used as headless attributes is exactly the same as when modifying an overt head; they must specify the number and noun class of their heads, whether overt or not. The same is true of verbs in relative clauses which are simply finite nominalized verbs, but the details of this are far too complex to go into here; see Foley (1991:413–433) for extensive discussion and analysis. Murik also has very few true basic adjectives; indeed *apo* 'big' may be the only clear member. Many property denoting words are derived from nouns, *aur* 'fire' > *aurarogo* 'warm', or verbs, *babaan* 'to be or act stupid' > *babago* 'stupid'. These can occur as modifiers with overt heads, *iranmot apak* (lit. houses.PL big.PL) 'big houses', or without, *apak* (lit. big.PL) 'big (things)', but in either case they must inflect for the number of their heads, overt or not.

2.3 Verbal number

This is not a feature of Yimas, nor, on available data, of the other three noun classifying languages. It was found in Murik as spoken in the nineteenth century and early twentieth century and documented by Schmidt (1953), but it is not known if it is still found in the contemporary language, and it is lacking in its closest relative, Kopar. Nineteenth century Murik is a very interesting case because, of course, it is one of the very rare languages worldwide which for nominal number distinguishes four numbers, singular, dual, paucal and plural, so how does verbal number work? Verbs that distinguish verbal number are common in the language and occur in two forms: one to indicate singular and dual number and the other to express paucal and plural number. The formal morphology that expresses this contrast is very complex and irregular and largely lexically specified, as verbs appear to fall into distinct inflectional classes for marking verbal number. It is beyond our discussion here, but its complexity can be grasped by the examples provided. For intransitive verbs verbal number indicates the number of the subject, *dəpraat* 'one or two fly away' versus *dəpraasan* 'three or more fly away', *dəkaraat* 'one or two get up' versus *dəkaraasan* 'three or more get up', *poanəmbageit* 'one or two fall down' versus *poanəmbagein* 'three or more fall down', whereas for transitive verbs it is the number of the object which is indicated, *gəgeit* 'break one or two things' versus *gəgein* 'break three or more things', *bagam* 'kill one or two people' versus *bəbaan* 'kill three or more people', *arin* 'put one or two things inside' versus *arisan* 'put three or more things inside', *etataam* 'take one or two things out' versus *etaran* 'take three or more things out'. Note that this binary distinction in verbal number contrasts sharply with pronominal number, which like nominal number exhibits the full four number contrasts (Table 1).

Verbal number is a distinct verbal morphological category and is different from the expression of verbal aspect, which is indicated by other verbal affixes. As far as the available data indicate, a plural verb can simply indicate multiple actors or objects acted upon within one action or it can express one or more participants over repeated actions (a so called pluractional or iterative verb). An example of an inherently plural verb that must be interpreted as pluractional is *aboatakeiam* 'hit someone's ears with one hand'; clearly with one hand one cannot hit more than one ear at the same time, so this plural verb must mean hitting ears repeatedly. A person only has two ears, so the meaning is hitting at most a dual object over three or more repetitive blows; the plural verb therefore expresses multiplicity of actions (pluractional), not objects acted upon (dual objects take singular/dual verb forms). While pluractionality seems an enforced choice in this example and probably in other examples like *bəbaan* 'kill three or more people', as it is hard to kill more than one person at a time with traditional weapons like an arrow or spear, so this requires repeated actions of shooting or spearing, or *panein* 'bend three or more bows', other examples are vague between multiple participants in one action or one or more participants in repeated actions, such as *dəpraasan* 'three or more fly away', *ansarin* 'stitch three or more things together', *sakopin* 'distribute three or more things' *tabogin* 'wash three or more things' and *takəb* 'cover, protect three or more things'.

3 Agreement and the syntax of number

All Lower Sepik languages that have obligatory number inflection on nouns require their modifiers including adjectives to agree in number with them (as Kopar lacks such obligatory nominal number marking, it also lacks such agreement), and the four noun classifying languages require class agreement as well. As Murik has lost its noun classes, adjectives and other modifiers like demonstratives agree only for number, but again for all four numbers:

(37)   adjective 'big'   demonstrative 'that'
  SG   *ap-o*   *dae-wa*
  DL   *ap-aabo*   *dae-bwa*
  PC   *ap-aara*   *dae-ra*
  PL   *ap-ak*   *dae-ga*

Possessors have an option, They can occur before their head noun without number agreement, *o-na ira-mbo* 2PL-POSS house-DL 'your houses', or they can follow and in this case must agree in number, *ira-mbo o-na-abo* (lit. house-DL 2PL-POSS-DL) 'your houses'. Adjectives always follow their head noun and must show inflection for number, but surprisingly allow an option where the number of the noun is only

marked on the modifier, *iran ap-aara* (lit. house big-PC 'big houses') and *bren ap-aara arat-aara* (lit. pig big-PC good-PC) 'good big pigs'; this is impossible in Yimas **numpran kpa-m* (lit. pig.III.SG big-III.PL), intended 'big pigs'.

An interesting and typologically unusual feature of Murik is number agreement on postpositions. Postpositions agree in number with their complements and this is in accord with the heavy head marking typology of the languages, although this seems a rare feature crosslinguistically. A good example is (38) (Schmidt 1953):

(38) *ma   nor-əmbo arat-aabo rəbo-nai tə-ma-o-ga*
    1SG man-DL   good-DL   DL-to   PERF-1SG.NOM-go-FUT
    'I will have gone to the two good men'

Sample number inflections of a few postpositions in Murik are (Schmidt 1953):

(39) | | 'to' | 'for' | 'with' |
|------|--------------|--------------|-------------|
| SG | *nai* | *ndo* | *ŋa* |
| DL | *rəbo-nai* | *rəbo-ndo* | *rəbo-ŋa* |
| PC | *gəta-nai* | *gəta-ndo* | *gəta-ŋa* |
| PL | *təmboa-nai* | *təmboa-ndo* | *təmboa-ŋa* |

For the noun classifying languages I will take Yimas and Chambri as typical and make use of additional data on Chambri from Pagotto (1976). All four noun classifying languages require all modifiers of a head noun to agree with it in number and class. There are distinct adjectival suffixes for each class and number specifications, as in these examples:

(40) a. Yimas
      yurantrm   kpa-ntrm
      dog.III.DL big-III.DL
      'two big dogs'

    b. Chambri
      upun-asim yuranasim
      big.I.DL   dog.I.DL
      'two big dogs'

The same suffixes as mark adjectival modifiers are also used on possessive pronouns, but Yimas allows the same alternatives as Murik: the possessive pronoun can occur in front of certain possessed nouns without the concordial suffix:

(41) *kra-na   ŋayuk*
    3PC-POSS mother
    'their mother'

But if postposed or floated away from the head noun, it is necessary:

(42) *ŋayuk   kra-na-nmaŋ*
     mother 3PC-POSS-II.SG
     'their mother'

The Chambri data available do not show this option; even if immediately before the head noun, the agreement suffix seems to be required:

(43) *mə-na-nəmp   mananəmp*
     3SG-POSS-I.DL husband.I.DL
     'her husbands'

Table 13 provides the set of adjectival/pronominal possessor agreement suffixes for all classes and numbers in Yimas and Chambri (recall Chambri only has seven classes while Yimas has ten; the classes are correlated etymologically in Table 13). The plural ending for Chambri in all classes is *-ar,* so it is the typologically rare case here of a language that makes more distinctions in the dual number than in the plural. For the few examples of nouns with human referents that inflect for paucal number, modifying adjectives also inflect for paucal number: *norakraŋk wupa-ŋk* man.PC big-PC 'big men'.

Tab. 13: Adjectival/possessor Agreement Suffixes in Yimas and Chambri.

Yimas			Chambri			Yimas			Chambri		
I	SG	*-n/-k*	I	SG	*-an*	VI	SG	*-ŋk*	IV	SG	*-aŋk*
	DL	*-mampan/-rm*		DL	*-anəmp*		DL	*-ŋkl*		DL	*-ŋkəkri*
	PL	*-ump*		PL	*-ar*		PL	*-ŋki*		PL	*-ar*
II	SG	*-nmaŋ*				VII	SG	*-mp*	V	SG	*-amp*
	DL	*-nprum*					DL	*-mpl*		DL	*-ampri*
	PL	*-nput*					PL	*-ra*		PL	*-ar*
III	SG	*-n/-kn*				VIII	SG	*-i*	VI	SG	*-ake*
	DL	*-ntrm*					DL	*-(i)l*		DL	*-riri*
	PL	*-ump*					PL	*-ra*		PL	*-ar*
IV	SG	*-um*	II	SG	*-m*	IX	SG	*-aw*	VII	SG	*-o*
	DL	*-mul*		DL	*-muri*		DL	*-awl*		DL	*-ori*
	PL	*-ra*		PL	*-ar*		PL	*-ut*		PL	*-ar*
V	SG	*-n/-kn*	III	SG	*-an*	X	SG	*-uŋk*			
	DL	*-ntrm*		DL	*-asim*		DL	*-uŋkl*			
	PL	*-ra*		PL	*-ar*		PL	*-ra/-uŋkwi*			

Demonstratives or deictic modifiers are more complicated in Yimas than in Chambri. Chambri is analyzed in Pagotto (1976) as having a binary distinction between a proximal demonstrative *maka-* 'this' and a distal one *kwa-* 'that', as well as third

person pronouns built on a stem in *m-*: SG: *m-ən* DL: *m-ənəmp* PL: *m-omp*. However, Yimas has no true third person pronouns; it uses demonstratives for this function and has three: proximal *-k* 'this (near me)', near distal *m-* 'that (near you)' (cognate to the Chambri third person pronoun base) and far distal *-n* 'that yonder near neither me nor you'. Only the near distal *m-* in Yimas occurs with the adjectival/possessive agreement suffixes of Table 13; the other two take the verbal agreement prefixes to be discussed below. In Chambri both demonstratives take agreement suffixes very similar to the adjectival/possessive agreement suffixes of Table 13 and clearly derived from them:

Tab. 14: Demonstrative Agreement Suffixes in Chambri.

I	SG	-ne	V	SG	-mpe
	DL	-nəmp		DL	-mpri
	PL	-mpwe		PL	-re
II	SG	-me	VI	SG	-ke
	DL	-muri		DL	-riri
	PL	-re		PL	-re
III	SG	-ne	VII	SG	-owe
	DL	-sim		DL	-ori
	PL	-re		PL	-re
IV	SG	-ŋke			
	DL	-ŋkəkri			
	PL	-re			

Number marking on verbs other than the verbal number of nineteenth century Murik discussed in section 2.3 is to signal the number of the verb's core arguments in portmanteau affixes with person. In these languages this is done mainly by prefixes, with the exception of Kopar, which mostly shifted to relying on suffixes, a development currently taking place in Kanda as well; examples from Kopar:

(44) a. *i-m-ar-aŋg-bako*
         2-eat-PROG-PRES-2DL
         'you are eating'

      b. *i-tumanəŋ-aŋg-oko*
         2-hit-PRES-2DL
         'you hit him/them'

All Lower Sepik languages are heavily head marking and lack nominal case marking for core arguments (though Kanda seems to be innovating an ergative postposition), so their roles are indicated by verbal cross-referencing affixes, up to three in the case of ditransitive verbs, as in this example form Yimas:

(45) *uraŋ   k-n-ŋa-r-mpun*
    coconut.VI.SG VI.SG.NOM-3SG.ERG-give-PERF-3PL.DAT
    'he/she gave the coconut to them'

The marking of roles is extremely complex in these languages, following a split ergative case system which is overlaid by a direct-inverse system (see Foley 2016). The system of cross-referencing prefixes for Murik is displayed in tables 6 and 7. Note that the first and second person singular forms show a three way split, the ergative, nominative and accusative are all distinct, but the non-singulars are simply nominative-accusative, the forms in ergative and nominative columns being the same. But the third person forms are always ergative-nominative, so that the nominative and accusative forms are identical. Note also the second and third singular ergative forms are both zero. This is a crosslinguistically rare but diagnostic property of Lower Sepik languages; these two forms are typically homophonous, even if not necessarily zero (in Yimas the form is *n-*), but again Kopar is an exception.

    These languages share a fairly widespread typological property of head marking languages, a direct-inverse inflectional system. In languages with a direct-inverse inflectional pattern, the pronominal agreement affixes are ranked according to their person, first and second persons, the local persons, outranking third person, the non-local person. When a higher ranked person is the actor and a lower ranked person the object of a transitive verb, that is the direct inflectional pattern; when the linking is reversed, that is the inverse inflection. In the Lower Sepik languages the system works according to the following rule: whichever pronominal agreement prefix refers to the higher ranked person occurs closer to the verb stem. But if both referents of both affixes are low ranked third persons, then that for the actor holds the higher rank and occupies that position, for example in Murik:

(46) *do-bo-kərə-na*
    3PC.NOM-3PL.ERG-hit-PRES
    'they hit them'

If one of the pronominal agreement prefixes is first or second person and the other third person, the former as the higher ranked always occupies the position next to the verb stem, regardless of its role as actor or object, and, when the higher ranked person is the object, the inverse situation, the Murik verb necessarily takes an inverse circumfix *n-...-ŋa*:

(47) a. direct
       *g-**a**-kərə-na*
       3PL.NOM-1SG.ERG-hit-PRES
       'I hit them'

b. inverse
*nu-mbu-**aŋa**-kərə-ŋa-na*
INV-3PL.ERG (INV)-1SG.ACC-hit-INV-PRES
'they hit me'

Note that the contrast between direct versus inverse realization does not alter grammatical functions, merely their morphological formal expression; this is not a diathetic voice contrast like active-passive.

As Murik and Kopar have lost the system of nominal classification, the bound verbal affixes only indicate person and number. The four noun classifying languages also indicate the noun class of third person core arguments. However, Chambri is depauperate here, only indicating noun class in the singular and then only for certain classes:

(48) a. *sraŋk   aŋkə-kəbran*
tooth.IV.SG IV.SG.NOM-break
'the tooth broke'

b. *anamp   ampə-kəbran*
bone.V.SG V.SG.NOM-break
'the bone broke'

c. *tənum   amu-kəbran*
sago_palm.II.SG II.SG.NOM-break
'the sago palm broke'

In the non-singular numbers there is no contrast for class:

(49) a. *sraŋkəkri   ari-kəbran*
tooth.IV.DL DL.NOM-break
'the two teeth broke'

b. *anampri   ari-kəbran*
bone.V.DL DL.NOM-break
'the two bones broke'

c. *sraŋkər   ar-kəbran*
tooth.V.PL PL.NOM-break
'the teeth broke'

d. *anabar   ar-kəbran*
bone.V.PL PL.NOM-break
'the bones broke'

For the other three noun classifying languages noun class is always indicated as well as number. But there is only one set of case affixes which signal the full range of noun class oppositions and that is the nominative. The ergative, accusative and dative cases never do, but as they are restricted essentially to human referents, i.e. nouns belonging to classes I and II, that is not surprising (because the case marking schema for third person participants is ergative-nominative, there can be no inanimate accusatives; any direct object noun with inanimate reference will be realized as nominative). Depending on mood and grammatical construction, noun class marking nominative affixes can be prefixes or suffixes, and the conditions determining which it will be may be very complex. Prefixes are generally the default realization in these languages, as in Yimas:

(50) a. *uraŋki*   *kia-tmuk-t*
        coconut.VI.PL VI.PL.NOM-fall-PERF
        'the coconuts fell'

    b. *uraŋki*   *kia-mpu-am-t*
        coconut.VI.PL VI.PL.NOM-3PL.ERG-buy-PERF
        'they ate coconuts'

In Kanda prefixes and suffixes seem about equally common, depending on transitivity and tense of verb:

(51) a. *waŋkli*   *ikolontə-ka-ŋkli*
        coconut.III.PL fall-PRES-III.PL.NOM
        'the coconuts are falling'

    b. *waŋkli*   *kle-m-amə-ŋ*
        coconut.III.PL III.PL.NOM-1SG.ERG-eat-PRES
        'I am eating coconuts'

Table 15 presents the nominative noun class marking verbal prefixes in Yimas, Karawari and Kanda (note that Karawari lacks a dual in some classes and Kanda lacks it entirely).

Of the three languages Yimas is the most conservative, and we will see more evidence for this below in discussing the numerals. Kanda has lost the dual number everywhere and Karawari in a number of classes. Clearly the Kanda plural morphemes for classes III, IV, and VI are cognate to the dual prefixes in the corresponding classes in Yimas. Karawari has also lost a number of other distinctions preserved in Yimas. For example, it has gone some way in the direction of Chambri by having fewer distinctions in the plural; *maya-* and *mi-* are spread over a number of classes. And there is no difference between the singular and dual prefix in class VIII. Surprisingly even though Yimas and Karawari are more closely related, in fact form a

Tab. 15: Nominative Noun Class Marking Verbal Prefixes.

Yimas			Karawari			Kanda		
I/II	SG	*na-*	I/II	SG	*mən-*	I	SG	*na-*
	DL	*impa-*		DL	*ŋkri-*			
	PL	*pu-*		PL	*mpu-*		PL	*pa-*
III	SG	*na-*	III	SG	*mən-*			
	DL	*tma-*						
	PL	*pu-*		PL	*maya-*			
IV	SG	*mu-*	IV	SG	*ma-*			
	DL	*mula-*						
	PL	*ya-*		PL	*mi-*			
V	SG	*na-*	V	SG	*mən-*	II	SG	*na-*
	DL	*-(nt)rm*						
	PL	*ya-*		PL	*maya-*		PL	*ta-*
VI	SG	*k-*	VI	SG	*ŋk-*	III	SG	*ke-*
	DL	*kla-*		DL	*ŋkri-*			
	PL	*kia-*		PL	*ŋki-*		PL	*kle-*
VII	SG	*p-*	VII	SG	*mp-*	IV	SG	*pa-*
	DL	*pla-*		DL	*mp-*			
	PL	*ya-*		PL	*maya-*		PL	*pela-*
VIII	SG	*i-*	VIII	SG	*mi-*	V	SG	*i-*
	DL	*ila-*						
	PL	*ya-*		PL	*mi-*		PL	*wa-*
IX	SG	*wa-*	IX	SG	*ma-*	VI	SG	*wa-*
	DL	*ula-*						
	PL	*ura-*		PL	*mi-*		PL	*wela-*
X	SG	*ku-*	X	SG	*ŋk-*			
	DL	*kula-*		DL	*ŋkri-*			
	PL	*ya-/kwia-*		PL	*maya-/ŋgi*			

sub-group, in these sets of nominative prefixes, those of Yimas and Kanda are more similar and more faithfully preserve the Proto-Lower Sepik forms. Finally, as mentioned above, the proximal demonstrative *-k* and far distal *-n* in Yimas (and equivalents in Karawari*)* take these verbal prefixes to agree with their modified nouns, not the adjectival/possessive suffixes used in Chambri:

(52) Yimas

    a. *uraŋki*   *kia-k*
       coconut.VI.PL VI.PL-PROX
       'these coconuts'

    b. *paypratawi* *pu-k*
       bird.III.PL   III.PL-PROX
       'these birds'

Both demonstratives and their head nouns must be inflected for number; inflecting one or the other is ungrammatical in Yimas:

(53) a. *uraŋki   k-k
     coconut.VI.PL VI.SG-PROX
     'these coconuts'

    b. *uraŋ   kia-k
     coconut.VI.SG VI.PL-PROX
     'these coconuts'

Other than the case of preposed possessors in Murik and Yimas illustrated above (and similarly the basic adjectives of Yimas), which do not require agreement, with the exception of Kopar, all modifiers of a head noun must agree in number with it, and in the noun classifying languages, in noun class as well. Word order within the NP can be rather flexible in these languages, although the degree of flexibility varies across languages; for example in Chambri adjectives can come either before or after the head noun (Pagotto 1976):

(54) a. *wupa-n dikek-an   yuri*
     big-I.SG black.I.SG dog.I.SG
     'a big black dog'

    b. *dikek-an wupa-n yuri*
    c. *yuri   wupa-n dikek-an*

However demonstratives must come before the noun, and numerals seem to favor this position also, though not obligatorily:

(55) *maka-sim wi-sim   amərəmpəkə-nasim yuranasim*
    PROX-I.DL two-I.DL hungry-I.DL   dog.I.DL
    'those two hungry dogs'

Murik is less flexible; modifiers including demonstratives always follow the head noun, but agree in number, even if the noun is not overtly marked for number:

(56) a. *bren ap-aara arat-aara*
     pig   big-PC   good PC
     'big good pigs'

    b. *ira(n)-moara ap-aara keroŋgo e-ra*
     house-PC   big-PC   three   PROX-PC
     'these three big houses'

Here are three examples of rather fully expanded noun phrases in Kanda:

(57) a. *imparŋkar ta   ami-na-klea   sum-erəm kup-le*
       pig.I.PL   PROX.I.PL 1SG-POSS-I.PL I.PL-three big-I.PL
       'these three big pigs of mine'

    b. *paruŋkli   kle   ami-na-ŋklea   kl-erəm   kup-aŋklea*
       betelnut.III.PL PROX.III.PL 1SG-POSS-III.PL III.PL-three big-III.PL
       'these three big betelnuts of mine'

    c. *səmur   wura   ami-na-kura   wa-rəm   kup-ura*
       cane.VI.PL PROX.VI.PL 1SG-POSS-VI.PL VI.PL-three   big-VI.PL
       'these three big (pieces of) cane of mine'

As with Chambri there is some flexibility in the ordering of the constituents in these Kanda noun phrases, though it is not clear just what the degree of flexibility is; surprisingly possessors follow here, whereas in Chambri they precede the head noun:

(58) a. *amə-n(a)-ar   numpranar*
       1SG-POSS-PL pig.I.PL
       'my pigs'

    b. *numpran mə(n)-n(a)-ar namaŋkər*
       pig.I.SG   3SG-POSS-PL leg.IV.PL
       'the pig's legs'

Nominal possessors follow in Murik:

(59) *ira(n)-mbo Got na-abo*
    house-DL   God POSS-DL
    'God's two houses'

Yimas is particularly interesting in this regard as it has very depauperate true noun phrases. True noun phrases can consist of at most two constituents, a head noun and a preposed possessor or adjective, and these with no agreement suffixes:

(60) a. *ama-na   matn*
       1SG-POSS brother.I.SG
       'my brother'

    b. *kpa patn*
       big betelnut.V.SG
       'big betelnut'

No permutations nor additional material are possible:

(61) a. *matn   ama-na*
    brother.I.SG 1SG-POSS
    'my brother'

    b. *ama-na kpa patn*
      1SG-POSS big betelnut.V.SG
      'my big betelnut'

Any more expansion requires agreement affixes on the modifiers and then the order is no longer fixed:

(62) a. *patn   ama-na-kn   kpan-n*
    betelnut.V.SG 1SG-POSS-V.SG big-V.SG
    'my big betelnut'

    b. *kpa-n   patn   ama-na-kn*
    c. *kpa-n   ama-na-kn   patn*
    d. *patn   kpa-n   ama-na-kn*
    e. *ama-na-kn   patn   kpa-n*

Of these permutations (d) and (e) are probably the most common. But it seems like these are not true noun phrases at all, just nominals in apposition to each other and strung together, for they can be floated away with ease:

(63) *patn   wayk-k   ama-na-kn   wa-n*
    betelnut.V.SG buy-PURP 1SG-POSS-V.SG go-IMP
    'go buy my betelnut'

Finally, it should be pointed out that there is no difference in morphology between attributive and predicative functions in these languages. Adjectives and other modifiers in both these functions must agree in number and, in the classifying languages, noun class, with the noun they modify. Yimas has a copula which also must show such agreement, but this is lacking in Murik:

(64) a. Murik
     daŋ o-n(a)-aara   pasig-aara
     foot 2PL-POSS-PC black-PC
     'your feet are black'

    b. Yimas
     m-n   kpa-n   anak
     NR.DIST-I.SG big-I.SG COP.I.SG
     'he is big'

The Yimas copula is optional though:

(65) *numpran   urkupwica-k-n*
      pig.III.SG blacken-STATIVE-III.SG
      'the pig is black'

Numerals in these languages are built on a base five, essentially counting fingers and then toes; 'ten' in Yimas is *nuŋkarawl* 'hand.IX.DL' 'two hands' and 'twenty' is *namarawt muntak-n* person.I.SG whole-I.SG 'whole person'. In the non-classifying languages, Murik and Kopar, numerals are invariable (though fossilized prefixes indicate they did not used to be), but in the others, numerals inflect to indicate the class of the noun enumerated. The number 'five' and its multiples 'ten' and 'twenty' are always invariable, but 'one' through 'four' and all numbers including them, meaning all numbers except the base 'five' and its multiples, must vary to indicate the class of the noun enumerated. The system is very complicated and was already being lost by younger speakers when I started researching these languages in the late 1970s – they just employed the simple Tok Pisin numerals – but it works essentially like this: 'one' inflects like an adjective, 'two' and 'three' like a verb, and 'four' can either inflect like an adjective as it does in Yimas or be invariable as it is in Karawari. 'One' is the most straightforward; it simply takes the corresponding adjectival endings for each class and number:

(66) a. Yimas
         wun   *mpa-n*
         sago_grub.V.SG   one-V.SG
         'one sago grub'

      b. Karawari
         səkər   *mpa-n*
         chair.V.SG   one-V.SG
         'one chair'

      c. Kanda
         wurən   *mpia-n*
         sago_grub.II.SG   one-II.SG
         'one sago grub'

      d. Chambri
         wun   *mbwia-n*
         sago_grub.III.SG one-III.SG
         'one sago grub'

(67) a. Yimas
trŋ   *mpa-ŋ*
tooth.VI.SG one-VI.SG
'one tooth'

b. Karawari
səsəŋ   *mpa-ŋ*
tooth.VI.SG one-VI.SG
'one tooth'

c. Kanda
sisiŋ   *mpia-ŋ*
tooth.III.SG one-III.SG
'one tooth'

d. Chambri
sraŋk   *mbwia-ŋ*
tooth.IV.SG one-IV.SG
'one tooth'

(68) a. Yimas
tanm   *mpa-m*
bone.VII.SG one-VII.SG
'one bone'

b. Karawari
tanəm   *mpa-m*
bone.VII.SG one-VII.SG
'one bone'

c. Kanda
mənəm   *mpia-m*
neck.IV.SG one-IV.SG
'one neck'

d. Chambri
anamp   *mbwia-m*
bone.V.SG one-V.SG
'one bone'

'Two' and 'three' are more complicated because there has been some reworking of the paradigms in the languages. Examples for 'two' follow:

(69) a. Yimas
wuntrm   *tm-pal*
sago_grub.V.DL V.DL-two
'two sago grubs'

b. Karawari
səkərŋkar səm-ipay
chair.V.PL V.DL-two
'two chairs'

c. Kanda
wurəŋkar   səm-upar
sago_grub.II.PL II.DL-two
'two sago grubs'

d. Chambri
wunasəm   wu-səm
sago_grub.III.DL two-III.DL
'two sago grubs'

(Kanda lacks the dual number entirely for nouns and Karawari for this class, but here preserves as a fossil the old dual concordial affix, still apparent on the head nouns and numerals in Chambri and Yimas)

(70) a. Yimas
trŋkl   k-rpal
tooth.VI.DL VI.SG-two
'two teeth'

b. Karawari
səsəŋkri   k-ripay
tooth.VI.DL VI.SG-two
'two teeth'

c. Kanda
sisiŋkli   kl-upar
tooth.III.PL III.PL-two
'two teeth'

d. Chambri
sraŋkəkri   k-ri
tooth.IV.DL IV.SG-two
'two teet

(71) a. Yimas
tanpl   p-rpal
bone.VII.DL VII.SG-two
'two bones'

b. Karawari
tanəmpri   p-ripay
bone.VII.DL VII.SG-two '
two bones'

c. Kanda
mənəmpar pl-upar
neck.IV.PL IV.PL-two
'two necks'

d. Chambri
anampri   p-ri
bone.V.DL V.SG-two'.
'two bones'

'Two' must always occur with a noun overtly inflected for dual number when that possibility exists; the option present in Murik of a numeral with a noun uninflected for number is not possible. Note in the first example for class V in Yimas and Karawari corresponding to class II in Kanda and class III Chambri, all four languages use a dual affix from Proto-Lower Sepik **səm-*, continued as a prefix generally but surfacing as a suffix in Chambri. For the other two sets, exhibiting classes VI and VII (Yimas/Karawari) and III and IV (in Kanda and IV and V in Chambri), all but Kanda use a singular verbal prefix as still preserved in the verbal paradigms of Yimas; Kanda has innovated and uses a plural prefix (formerly, a dual, as for these classes in Kanda the synchronic plural marker is the older dual formative).

Examples for 'three' follow:

(72) a. Yimas
wunt   tamunum
sago_grub.V.PL three.V
'three sago grubs'

b. Karawari
səkərŋkar səm-ianmaw
chair.V.PL V.DL-three
'three chairs'

c. Kanda
wurəŋkar   səm-eləm
sago_grub.II.PL II.DL-three
'three sago grubs'

d. Chambri
wunar   samnenamp
sago_grub.III.PL three.III
'three sago grubs'

(73) a. Yimas
trŋki   k-ramnawt
tooth.VI.PL VI.SG-three
'three teeth'

    b. Karawari
       səsəŋki   k-rianmaw
       tooth.VI.PL VI.SG-three
       'three teeth'

    c. Kanda
       sisiŋkli   kl-eləm
       tooth.III.PL III.PL-three
       'three teeth'

    d. Chambri
       sraŋkər   kia-ram
       tooth.IV.PL IV.PL-three
       'three teeth'

(74) a. Yimas
       tanpat   p-ramnawt
       bone.VII.PL VII.SG-three
       'three bones'

    b. Karawari
       tanəmpas   p-rianmaw
       bone.VII.PL VII.SG-three
       'three bones'

    c. Kanda
       mənəmpar pl-eləm
       neck.IV.PL IV.PL-three '
       three necks'

    d. Chambri
       anabar   pia-ram
       bone.V.PL V.PL-three
       'three bones'

'Three' must always occur with a noun overtly inflected for plural number (or paucal number for the small set of overtly inflected paucal nouns in Chambri); again the Murik option of a noun uninflected for number is not possible. Note in the first set of examples classes V/II/III, Yimas and Chambri have a cognate suppletive form for 'three'. This is probably the original situation, and Karawari and Kanda have spread the base of the numeral from the other classes to this. For the other two sets of examples, classes VI/III/IV and VII/IV/V, Yimas and Karawari again use the singular prefix for noun class agreement, but Kanda and Chambri use a plural one. What is the original situation? The clue is in the now invariable numerals of Murik and Kopar, *keroŋgo* and *kereməŋ* respectively, which have the fossilized singular

prefix *k-* as in Yimas and Karawari for class VI. Kanda has simply done with 'three' what it did with two', i.e. use its plural prefix, again, of course, formerly a dual before Kanda lost the dual number. Chambri is more interesting. While it uses a singular prefix for 'two', it uses a plural prefix for 'three', but these prefixes *kia-* and *pia-* exist nowhere else in the contemporary language. The scenario seems to have been that Chambri has innovated by using plural prefixes in these classes, but not elsewhere; for example, *tənil mu-ramtam* sago_palm II.SG-three 'three sago palms', but these prefixes then disappeared elsewhere in the language because Chambri does not indicate noun class in its verbal cross-referencing prefixes for plural, it only marks number. But it must have done so in earlier stages because one of these is directly cognate with Yimas *kia-,* the nominative verbal prefix for class VI plural (see Table 15) and it must have been from the verbal system of affixation that these plural prefixes spread to 'three'.

The numeral 'four' like 'one' in Yimas also inflects like an adjective, but in an unusual way. The base of 'four' is the adjective *ma-* 'other', and this is reduplicated, to which is added the dual suffix appropriate to the class of the noun being enumerated, hence it means something like 'two twos':

(75) a. *wunt*   *ma-r(m)-ama-rm*
       sago_grub.V.PL other-V.DL-other-V.DL
       'four sago grubs'

    b. *tanpat*   *ma-mpl-ama-mpl*
       bone.VII.PL other-VII.DL-other-VII.DL
       'four bones'

    c. *wunamarut*   *ma-(a)wl-ama-(a)wl*
       centipede.IX.PL other-IX.DL-other-IX.DL
       'four centipedes'

There are three quantifiers in Yimas *muntak* 'many, all', *mpatnpa-* 'a few' and *ntuknti* 'how many, how much'; they only occur with plural nouns. *Muntak* 'many' is invariable: *wunt muntak* (lit. sago_grub.V.PL many) 'many sago grubs', but *mpatnpa-* '(a) few' varies for class: *tanpat mpatnpa-ra* (lit. bone.VII.PL few-VII.PL) 'a few bones'

Numerals appear to have rather fixed ordering in these languages; with the exception of Chambri, which seems to permit both preposed and postposed positions, they always occur after the enumerated noun, though not necessarily immediately after (see Kanda examples in (57)). Even in Yimas, with its flexible and depauperate noun phrase structure, numerals follow the noun, but do not seem to form a constituent with the noun because they need not be adjacent to it. Numerals are strictly attributive. They are never predicative as in Austronesian languages; a sentence like 'the books were three' is impossible in these languages and rather one

must say 'there were three books'. Attributive numerals can occur without an overt enumerated noun as long as that is established in context, and again in the noun classifying languages, the numeral must indicate the number and noun class of its missing head, as in this Yimas question-answer pair:

(76) Q: *ipwa awŋkwi ntuknti kantk-um aypwak?*
    2PL   egg.X.PL how_many with-PL   COP.2PL
    'how many eggs do you have?'

  A: *ku-ramnawt*
    X.SG-three
    'three'

(Note the interesting grammatical feature of this construction in which the possessive postposition *kantk-* 'with' must agree in number (and noun class if relevant) with the possessor). When nouns are coordinated, the number indicated by modifiers and cross-referencing verbal affixes is their sum total, as these examples show:

(77) a. Chambri (Pagotto 1976)
    mbwia-n sinemp nəme san manan empu-wa-i
    one-III.SG day.III.SG wife.II.SG and husband.I.SG 3DL.NOM-go-PAST
    'one day a wife and husband went'

  b. Yimas
    Aympt Barati m-rm tan impa-nanaŋ-taw-ntut
    PN   PN   NR.DIST-II.DL there 3DL.NOM-DUR-sit-REM.PAST
    'Aympt, Barati, those two were living there'

  c. Murik (Schmidt 1953)
    məŋgə kiap graŋa də-ya-ra
    3PC   patrol_officer together 3PC.NOM-come-PAST
    'they and the patrol officer came together'

This summation of number expressed by the bound pronominal affixes is also how inclusory plurals are expressed in Yimas:

(78) *Yakayapan manpa na-ŋkra-tu-t*
    PN   crocodile.III.SG III.SG.NOM-1DL.ERG-kill-PERF
    'Yakayapan and I killed the crocodile'

The dual marking of the first person ergative pronominal indicates that the actor of the clause is the speaker including someone else, here the mentioned person called Yakayapan.

4 Semantics and discourse

The independent pronouns in Lower Sepik languages are not used to indicate politeness or honorification, as plural pronouns are used in the classical T/V pattern of European languages. These are extreme head marking languages, and independent pronouns are infrequent and carry strong contrastive emphasis. Rather it is in the complex verbal inflectional systems of Murik, Kopar and Karawari that we find a somewhat parallel phenomenon at work. Here I will illustrate with data from Kopar; Murik is similar, though slightly more elaborate. For third person subjects with intransitive verbs, Kopar distinguishes four numbers, typically marked by both prefixes and suffixes:

(79) a. *u-m-ar-aŋ-oya*
    3SG-eat-PROG-PRES-3SG
    'he is eating'

   b. *mbi-m-ar-aŋ-baya*
    3DL-eat-PROG-PRES-3DL
    'they both are eating'

   c. *ŋgi-m-ar-aŋ-iya*
    3PC/PL-eat-PROG-PRES-3PC
    'they few are eating'

   d. *ŋgi-m-ar-aŋ-gaya*
    3PC/PL-eat-PROG-PRES-3PL
    'they many are eating'

For transitive verbs suffixes now are the sole indicators of the number of the subject, but in addition it is expressed with a prefix, only the third plural *mbu-* functioning as an ergative prefix, but with all number contrasts in this prefix neutralized:

(80) a. *mbu-tumanəŋ-aŋ-oya*
    3.ERG-hit-PRES-3SG
    'he hit him/them'

   b. *mbu-tumanəŋ-aŋ-odudu*
    3.ERG-hit-PRES-3PC
    'they few hit him/them'

The only way here to indicate the number of the object in these sentences is to use independent pronouns, but that is rare and would express a contrastive focus, i. e 'it was him that they two hit'; normally context is sufficient to disambiguate. In a corresponding direct verb with a first or second person subject acting on a third

person object, the same holds: the person-number distinctions for the subject are preserved, and again expressed by suffixes, but are lost for the object:

(81) *i-tumanəŋ-aŋg-ona*
       2-hit-PRES-2SG
       'you hit him/them'

When a first person actor acts upon a second person object in transitive verbs, a strategy called impersonalization occurs, parallel to the use of third person plural *Sie* for a polite second person in German. Here the first person is realized by the third person plural ergative prefix *mbu-*, with all number contrasts neutralized, but now the suffix indicates the number of the second person object:

(82) a. *mbu-tumanəŋ-aŋg-bako*
          3.ERG (=1)-hit-PRES-2DL
          'I/we hit you two'

      b. *mbu-tumanəŋ-aŋg-aya*
         3.ERG (=1)-hit-PRES-2SG
         'I/we hit you'

Note that due to impersonalization number distinctions here are lost for the first person subject, but preserved for the second person object, in contrast to the examples of (80) and (81). Murik differs from Kopar in that impersonalization occurs whenever a local person, first or second, acts upon another local person:

(83) *nu-mbu-ŋe-kərə-ŋa-na*
       INV-3nSG.ERG(=2)-1SG.DL.ACC-hit-INV-PRES
       'you (or they) hit us two'

(see Foley (2016) for more detail on the feature of impersonalization and number neutralization in both Murik and Kopar).

Karawari also has an interesting wrinkle here. In this language the second plural accusative prefix is *i-:*

(84) *i-mp-i-kra-r*
       INV-3nSG.ERG-2PL.ACC-hit-PERF
       'they hit you'

But in a parallel to V phenomenon in which the plural indicates politeness or formality, this second person plural accusative prefix has generalized to indicate all inverse forms, i.e. when a lower ranked person acts on a higher ranked, and including all local persons acting on local combinations:

(85) a. *i-mp-ŋa-kra-r*
   INV-3nSG.ERG-1SG.ACC-hit-PERF
   'they hit me'

   b. *i-Ø-ŋa-kra-r*
   INV-2SG.ERG-1SG.ACC-hit-PERF
   'you hit me'

And somewhat like Kopar, if the actor is first person non-singular and the object second person, there is impersonalization to third:

(86) a. *i-mp-i-kra-r*
   INV-3nSG.ERG (=1nSG)-2PL.ACC-hit-PERF
   'we hit you'

   b. *i-mp-ŋki-kra-r*
   INV-3nSG.ERG (=1nSG)-2DL.ACC-hit-PERF
   'we hit you two'

In these languages, the third person plural is used as a generic pronoun like English *one* or *you*. Obviously, this has favored the process of impersonalization discussed above. In generic sentences normally the singular form of noun is used, though there is one example of a plural noun being used generically in Yimas, *tpwi* 'sago', the staple food, but the dual is never possible, as it always has specific reference. Here is a generic sentence from Murik (Schmidt 1953); note the subject is singular, but the corresponding English sentence allows either singular or plural:

(87) *majen   arəkəmə daat-agu-rogo*
   child.SG walk   know-NEG-SG
   'A small child cannot walk/small children cannot walk.'

Morphologically plural nouns generally correspond to count nouns in Lower Sepik languages, but there is a subset of them that have mass denotation. Examples in Yimas include *yat* 'blood', *ɲct* 'urine', *wapwi* 'hair', *awt* 'fire', *yaki* 'tobacco', *amtra* 'food', *maramara* 'cargo, things, stuff' and *awi* 'lime powder', and in Karawari (Telban 1992), *makwi* 'sap', *məndi* 'feces' (the equivalent *mlm* in Yimas is singular), *səndi* 'urine', *pari* 'spit' and *maray* 'water'. Some nouns, particularly foodstuffs, can be used as either mass or count. Like English nouns such as *coffee,* there are some nouns which when dual or plural are always count, but when singular can be either mass or count, Yimas *patn* 'betelnut, a betelnut' but *parŋkat* 'betelnuts'. All Lower Sepik languages lack determiners, and there is no formal marking of count versus mass nouns corresponding to English *a* versus *some* or *many* versus *much;* for exam-

ple, the Yimas quantity interrogative *ntuknti* appears with both count and mass nouns, i.e. there is no formal distinction for 'how much' versus 'how many'.

Number and noun class play the key role in tracking discourse referents in Lower Sepik languages. As heavily head marking languages they exhibit extensive ellipsis of arguments whose referents have already been established in context or cotext. Clauses commonly consist of just the verb or at most a verb and one associated noun phrase, so the marking of core arguments by verbal cross-referencing affixes is the major reference tracking device in these languages. Once a referent is established by an overt noun or noun phrase, it usually is not mentioned again, but simply tracked by verbal agreement affixes, and this can carry on for many clauses. Crucially it is the full system of number distinctions and in the noun classifying languages, the rich distinctions of class, that, in combination with the nominative-accusative and ergative-nominative case system, generate such a wide array of differentiated verbal affixes that allows referents to be preserved without overt nominal mention over such long stretches of text. For example, in the Yimas text *The Flood* (Foley 1991:477–482), the two protagonists, two sisters, Yapalmay and Mampalmay, are introduced in the first clause. For the nearly all of the rest of the text of some twenty-eight sentences, some quite long containing multiple clauses, they are simply monitored by the verbal affixes *impa-* 3DL.NOM and *mp-* 3DL.ERG.

5 Conclusion

Number is a pervasive and elaborated grammatical category in Lower Sepik languages, perhaps even on a scale that few other language families can match. While other languages, particularly Oceanic languages (Lynch, Ross and Crowley 2002) can match them in their number distinctions in pronouns and perhaps even exceed them, as with Mussau (Ross 2002), no other language to my knowledge besides Murik inflects nouns and their modifiers so thoroughly for four number distinctions, and few language families mark four numbers in verbal cross-referencing affixes for core arguments as most of the Lower Sepik languages do. The number category paucal is relatively rare among the world's languages, and the scale of its usage in the Lower Sepik languages has few if any peers elsewhere.

The following table summarizes the scope of number distinctions across nouns, independent pronouns and bound pronominals in the six languages (parentheses indicate optional or partial marking (i.e. dual and paucal only in first person in Kanda or paucal only in first person Chambri pronouns, or only a few nouns with paucal marking in Chambri, etc)):

Tab. 16: Summary of Number Distinctions across Lower Sepik Languages.

	Nouns	Independent Pros	Bound Pros
Murik	SG/DL/PC/PL	SG/DL/PC/PL	SG/DL/PC/PL
Kopar	(SG/PL)	SG/DL/PC/PL	SG/DL/PC/PL
Yimas	SG/DL/PL	SG/DL/PC/PL	SG/DL/(PC)/PL
Karawari	SG/(DL/PC)/PL	SG/DL/PL	SG/DL/PL
Kanda	SG/PL	SG/(DL/PC)/PL	SG/(DL/PC)/PL
Chambri	SG/DL/(PC)/PL	SG/DL/(PC)/PL	SG/DL/PL

Unlike some other languages animacy does not play any role in determining whether a noun will be inflected for number in the four noun classifying languages; any count noun regardless of the animacy of its referent must be inflected for number. However, in Kopar with optional number marking, the animacy hierarchy is relevant: nouns with human referents commonly are marked for number, whereas those with non-human referents very rarely are. And in Chambri, while number is a ubiquitous and obligatory inflectional category for count nouns, only those with human referents have unique paucal forms; nouns with non-human referents have paucals derived from base plurals. Animacy commonly plays an important role in pronominal systems, especially in bound pronominals, in all Lower Sepik languages. In Kanda only first person pronouns, the highest position on the animacy hierarchy of speaker > addressee > human > animate > inanimate, exhibit the full range of number contrasts, singular, dual, paucal and plural, while in Kopar and Murik when a participant of higher animacy acts on one of lower animacy we get direct verb inflection, but in the opposite case, inverse inflection. And quite notably in Yimas only participants with human referents display the full range of case oppositions in bound pronominals, ergative-nominative-accusative-dative, while those of lower animacy or inanimate reference appear in a single case form, the nominative.

Number inflection and conceptual quantitativity have a close correlation in Lower Sepik languages. An explicit dual marking always means 'two' exactly, paucal 'a few' and plural 'more than a few' (in languages that have not lost the paucal). The only exceptions are some mass nouns that have invariable plural inflection, although most mass nouns do occur in invariable singular forms. Dual number for count nouns is used equally with and does not distinguish between pairs of objects and a simply enumerated total of two objects: Yimas *tuŋkuruŋkl* eye.VI.DL 'two eyes' and *trŋkl* tooth.VI.DL 'two teeth'.

Number marking in bound pronominals is pervasive and carries a very high functional load in discourse in these languages. Given the very high rate of ellipsis of nouns with established referents, number marking is crucial to maintaining referents across discourse. Number agreement is spread throughout the noun phrase, as in the typical examples from Kanda in (57), and the number of the person and number of its core arguments are always marked on the verb, up to three in the case of

ditransitive verbs. Consequently, the grammatical machinery of these languages is very elaborated to signal number. Essentially, the two major parts of speech show a sharp split in how number is signaled, nouns and their modifiers like adjectives and demonstratives by suffixes and verbs largely by prefixes. Murik goes even further and marks postpositions by prefixes for the number of their complements.

There are many areas in these languages that need further research. Only one of them, Yimas, has been studied in some depth, and all six are very endangered, Kopar and Yimas being moribund. All of them should be studied much more, and particularly Kopar, given its very parlous state and very unusual verbal inflectional system. On the basis of my short period of a month's fieldwork I have a sketch grammar of Kopar in press, but this fascinating and complex language calls out for a detailed study before it is lost forever. Although I possess reasonably extensive field notes and will publish more on Kanda in due course, it also is a very intricate language which calls out for further study. It would also be very valuable to study the paucal number in detail in Murik, Chambri, Karawari and Kanda, as little is known about its usage and distribution in these languages, and paucal number is rare in the world's languages, so we should obtain as much information about it from as many sources as we can while this is still possible. Contemporary Murik also calls out for further study to compare it with its form from a hundred years ago documented by Schmidt (1953) and to see how the category of number has developed over that time. The category of verbal number is only attested in Murik and that in its nineteenth century and early twentieth century form; it would be valuable to know its status in the contemporary language, and if present in any of the other languages, though it is clearly lacking in Yimas and seems to be so in Kopar as well. Finally, generic sentences are very little researched in any of these languages and should be; even in Yimas the data on them are paltry.

Abbreviations

1	first person
2	second person
3	third person
I–X	noun classes
ACC	accusative
AUX	auxiliary
COP	copula
DAT	dative
DL	dual
DUR	durative
ERG	ergative
FUT	future tense
HAB	habitual
INV	inverse

NEG	negative
NFN	non-finite
NOM	nominative
NR.DIST	near distal
nSG	non-singular
PC	paucal
PERF	perfective
PL	plural
POSS	possessor
PRES	present tense
PROG	progressive aspect
PROX	proximal
REM	remote
SG	singular

References

Foley, William. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.

Foley, William. 1991. *The Yimas language of New Guinea.* Stanford: Stanford University Press.

Foley, William. 2005. Linguistic prehistory in the Sepik-Ramu basin. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan pasts* 109–144. Canberra: Pacific Linguistics.

Foley, William. 2016. Direct versus inverse in Murik-Kopar. In Jens Fleischhauer, Anja Latrouite & Rainer Osswald (eds.), *Explorations of the syntax-semantics interface*, 265–288. Düsseldorf: Düsseldorf University Press.

Foley, William. 2017a. The languages of the Sepik-Ramu basin and environs. In Bill Palmer, (ed.), *The languages and linguistics of the New Guinea area*, 177–411. Berlin: de Gruyter-Mouton.

Foley, William. 2017b. The polysynthetic profile of Yimas, a language of New Guinea. In Michael Fortescue, Marianne Mithun & Nickolas Evans (eds.), *The Oxford handbook of polysynthesis,* 808–829. Oxford: Oxford University Press.

Hale, Kenneth. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1. 5–47.

Lynch, John, Malcolm Ross & Terry Crowley (eds.), 2002. *The Oceanic languages.* Richmond, UK: Curzon Press.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62. 56–119.

Pagotto, Louise. 1976. The noun phrase in Chambri. *Department of Language University of Papua New Guinea Occasional Papers* 5. Port Moresby: University of Papua New Guinea.

Ross, Malcolm. 2002. Mussau. In John Lynch, Malcolm Ross & Terry Crowley (eds.), *The Oceanic languages,* 148–166. Richmond, UK: Curzon Press.

Schmidt, Joseph. 1953. *Vokabular und Grammatik der Murik-sprache in nordost-Neuguinea* (Micro-Bibliotheca Anthropos 3). Posieux: Anthropos Institute.

Telban, Borut. 1992. The Grammar of Karawari, East Sepik Province, Papua New Guinea. Unpublished Manuscript, Department of Anthropology and Prehistory, Faculty of Arts, Australian National University.

Wayan Arka and Mary Dalrymple

16 Number in Marori

Abstract: We describe the nominal and verbal number systems of Marori, a Trans-New Guinea/Papuan language. The expression of number in Marori is distinguished by distributed underspecified exponence. Certain number values can be expressed even when there is no dedicated number marking for that value: for example, third person dual nominal number is expressed not by dual morphology, but by a combination of nonsingular and nonplural exponents. Verbal number, including duactionality, is also characterized by distributed and underspecified exponence, with duactionality expressed by a combination of nonsingular and nonplural verbal number marking. Marori also allows specification of incompatible values for number, with particular semantic effects; a singular argument indexed by nonsingular verb marking gives rise to a comitative-inclusory interpretation, for example. These characteristics make the nominal and verbal number systems of Marori very different from the familiar syntactic number systems of languages like English.

1 Overview

Marori (also known as Morori/Moraori; ISO 639–3: mok) is a subgroup-level isolate, Trans-New Guinea/Papuan language, highly endangered, with around 16 fluent speakers in the village of Wasur, Merauke, Indonesia. Like other Papuan languages in Southern New Guinea (Evans et al., 2017), Marori is an inflectional language, showing a complex three-way number system, encoded by contrasting pronominal forms and complex verbal morphology.

The Marori nominal number system encodes a three-way distinction: singular (SG), dual (DU) and plural (three or more, PL) (Arka, 2011, 2012a,b, 2015). However, the three-way SG/DU/PL distinction is explicitly marked only for first and second person bound pronominals on the verb, and only for the A (actor) argument. Singular vs. nonsingular (more than one, NSG) or plural vs. nonplural (one or two, NPL) distinctions are relevant elsewhere, including for the third person bound/free pronominals. There are also general (GEN) number forms, which are unspecified for number. The distinctions encoded in the Marori number system are depicted in the chart in (1):

(1)

1	2	3	4 ...
GEN (any number)			
SG	NSG (more than one)		
NPL		PL (more than two)	

Common nouns in Marori have no number marking. However, there are several nouns which are lexically specified for SG or NSG number, such as *parapur* 'girl.SG'/ *moipur* 'boy.SG' (in which the morpheme *pur* means 'small/young'), *meninggon* 'children.NSG', *nembin* 'parent/son/daugher-in-law.SG', *nemnembin* 'parents/sons/ daughters.in.law.NSG', *ai* 'aunt.SG', *ayon* 'aunt.NSG', *bab* 'uncle.SG', *babon* 'uncle.NSG', *yoropur* 'elderly.person.SG', and *kofepurpur* 'elderly.person.NSG'. Derived predicative adjectival nominals are inflected for a SG vs. NSG distinction. Determiners also mark number within the noun phrase when they are present, distinguishing SG from NSG number.

Marori verbs encode pluractionality, or verbal number. Verbal number interacts with nominal number in a complex way, making use of overlapping resources. Verbal number marking can indicate participant plurality, event plurality, or both. In Marori, this depends on the lexical aspectual class of the predicate.

Like other languages of Southern New Guinea, Marori shows distributed underspecified exponence in marking. The NUM feature is distributed across different sites within the verb (and also across words). This allows number to be expressed without a dedicated number marker for each number value, either in morphology or in syntax.

1.1 Nominal number in Marori

As noted above, the Marori nominal number system encodes a three-way distinction: singular (SG), dual (DU) and plural (three or more, PL). Formal morphological coding is often underspecified as nonsingular (two or more, NSG) or nonplural (one or two, NPL), as shown in (2).

(2)	Category:	Formal number distinction:	Ref.:
	a. bound pronouns on the verb	3-way distinction, 1/2A: SG vs. DU vs. PL	Tables (10) and (11)
		2-way distinction, 3A: NPL vs. PL	
		2-way distinction, 1/2U: SG vs. NSG	
	b. free pronouns	2-way distinction: SG vs. NSG	Table (19)
	c. demonstrative and spatial deictics	2-way distinction: SG vs. NSG	Figure (22)
	d. common nouns	general number, other than a very few nouns	example (25)
	e. derived event nominal	2-way distinction: SG vs. NSG	example (26)

As shown in (2), the distribution of number marking follows the Animacy Hierarchy (Silverstein 1976; Dixon 1979; Corbett 2000, Chapter 3), with a threeway SG/DU/PL distinction applying only in the top segment of the hierarchy (first and second person bound pronominals on the verb), and only for the A (Actor). The two-way SG/NSG or PL/NPL distinctions are relevant elsewhere, including for the third person free pronominals. Note, however, that number marking for third person bound pronominals on verbs is expressed by means of a portmanteau morpheme that also expresses gender information; see Section 1.3.

Marori has a two-gender system (masculine/feminine) based on semantic (natural) gender rather than grammatical gender. Natural gender encompasses gender of animals that have clear and easily identified sex, such as pigs, kangaroos, and dogs. Other animates without (clear) natural sex, such as 'fish', and inanimates such as 'house' are referred to by the masculine form when they are singular. Nouns with plural reference are not distinguished in terms of their gender; they show the same agreement pattern as singular feminine entities. Gender is encoded for third person singular arguments of verbs by means of vowel alternations, as we discuss in Section 1.3. There are no classifiers.

1.2 Verbal morphology

The simplified template for inflected verbs in Marori is shown in (3).

(3)

	VERBAL NUMBER		
AFF^{-1}	ROOT	(AFF^{+1})	AFF^{+2}
(PERS/NUM)	(PERS/NUM/GEND)	(PERS/NUM/ASP)	(PERS/NUM/TENSE/ASP/MOOD)
U		A	

Referential (number and person) features of core arguments A (Actor) and U (Undergoer) are typically required to be indexed on the verb (an exception is inclusory constructions, as we discuss in section 2.6). Besides bearing argument indexing, verbs are the only lexical category which is inflected for tense, aspect and modality (TAM). Pluractionality, or verbal number is marked by root suppletion, suffixation to the root (i.e. AFF^{+1}), or both. Verbal forms that are not marked for pluractionality are nonpluractional, and do not refer to a multiplicity of events.

Finite verbs are typically auxiliary or light verbs, e.g. *yunggoru* in (4-a). However, certain non-auxiliary verbs of high frequency, such as *kundo* 'run' (4-b), are directly affixed with TAM morphology. In contrast, nouns in general (including event nouns, such as *kamae* 'anger' in (4-a)) are typically bare. Certain property nouns are derived showing a number distinction; see section 2.3. As shown in (4-a), the event noun functioning as a lexical predicate precedes the inflected auxiliary verb.

(4) a. pa=na   kamae yu-nggo-ru.
       soon=1SG   anger   1SG.U-AUX.NPL-1SG.A.FUT
       'I'll soon be angry (lit. I will soon undergo anger).'

    b. mbe=na   kundo-ru
       MBE=1SG   run-1SG.A.FUT
       'I will run.'

Core A and U argument roles are expressed by verbal indexing showing number inflection. The free U NP can be flagged by a case enclitic =*i*, which shows no number marking, as further discussed in Section 1.4.

1.3 Argument indexing

As shown in the inflectional verb template in (3), argument indexing on the verb encodes the U (prefix) and A (suffix) arguments separately.[1] Like the free pronouns, the U prefixes in (5) show a two-way (SG vs. NSG) number distinction. They differ in that there are no overt third person forms: we represent this as a null prefix Ø–. The exponent of the U prefix for the second person, *k-*, seems to have originated from the corresponding free pronoun.

(5) Undergoer prefixes in Marori:

	1	2	3
SG	i-	k-	Ø-
NSG	iar-	kar-	Ø-

When the U argument is third person singular, the root is also inflected for gender, with front and non-low vowels (*e* or *i*) marking '3M.SG' and back and low vowels (*a* or *o*) indicating '3F.SG/3NSG', and potentially also indicating TENSE depending on the verb.[2] Vowel harmony spreads across the auxiliary and the lexical predicate within the phonological word, as exemplified in (6); the harmony includes front vs. low or back qualities.[3]

---

1 See Haspelmath (2019) for a discussion of the concept of indexing/flagging and its relation to the widely adopted concept of "head/dependent" marking.

2 For non-third person, the vowel *a* or *o* is used, as seen in the following example with a first person U argument:

(i)   Thomas na=i   ter=i-mo-f
     Thomas 1SG=U   hit.NPL=1.SG.U-AUX-3NrPST
     'Thomas hit me.'

3 In (6), we treat the lexical predicate *ter/tor* 'hit.NPL' as a proclitic on the auxiliary *me-ben*; further research is needed on the details of prosodic structure in Marori.

(6) a. ter=Ø-me-ben
   hit.NPL=3U-3M.SG.U.AUX-1NPL.A.NrPST.PFV
   'I hit him (once or twice).'

  b. tor=Ø-mo-bon
   hit.NPL=3U-3F.SG.U.AUX-1NPL.A.NrPST.PFV
   'I hit her (once or twice).'

These features are realized on the light or auxiliary verb, which comes with its own underlying underspecified vowel: for example, the surface AUX forms [*mi*]~[*ma*] 'NFUT' and [*me*]~[*mo*] 'FUT' are potential realizations of the underlying template *mV* for the AUX. The U vowel marking (i.e. the height and backness features) expressing the U's gender spreads through the AUX and the lexical predicate. Examples (7)–(8) illustrate a four-way distinction, with [*e*] for masculine future and [*i*] for masculine non-future tense. The vowel expressing feminine singular, future tense is [*o*], and feminine plural or feminine non-future is expressed as [*a*]. The full set of contrasts for *ma* is shown in (9). Other verbs may show fewer distinctions, with [*a*] contrasting with [*e*] '3SG.M'/[*o*] '3SG.F/3NSG', irrespective of TENSE, as for the verb *ifo* 'see' in example (25). Vowel harmony typically spreads to the whole phonological word in PAST paradigms as in (7), whereas in present and future paradigms the first person A suffixes (e.g. *-ru* '1SG.A.FUT' and *-ren* '1DU.A.FUT') carry an invariant vowel and are not involved in vowel harmony, as shown in (8).

(7) Past paradigm of the root *kswe-* 'hit.PL'
  a. keswe=mi-men
   hit.PL=3M.SG.U.AUX-1NPL.A.NrPST.IPFV
   'I hit him (several times/repetitively).'

  b. kaswa=ma-mon
   hit.PL=3F.SG.U.AUX-1NPL.A.NrPST.IPFV
   'I hit her/them (several times/repetitively).'

(8) FUT paradigm of the root *kswe-* 'hit.PL'
  a. keswe=me-ru
   hit.PL=3M.SG.U.AUX-1SG.A.FUT
   'I will hit him (several times/repetitively).'

  b. koswo=mo-ru
   hit.PL=3F.SG.U.AUX-1SG.A.FUT
   'I will hit her (several times/repetitively).'

  c. keswe=me-ren
   hit.PL=3M.SG.U.AUX-1DU.A.FUT
   'We two will hit him (several times/repetitively).'

    d. koswo=mo-ren
       hit.PL=3F.SG.U.AUX-1DU.A.FUT
       'We two will hit her (several times/repetitively).'

(9)   Allomorphic space of the AUX *ma*

	SG		DU/PL
	3M	**3F**	**(no GEND distinctions)**
RPST/NrPST/PRES	mi		ma
FUT	me	mo	

The A affixes (Tables (10) and (11)) show a two-way or three-way distinction, depending on the person category. All A affixes clearly differ formally in their shapes from the U prefixes and from their free pronoun counterparts. They also show varying patterns of syncretism, indicated by differences in shading of the cells.[4]

(10) Class 1 Actor affixes in Maori:

	IRR/FUT			NrPST (Perfective)			RmPST (Perfective)		
	1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
SG	-ru	Ø	-Ø	-ben	-f	-f	-fori	-fi	-fi
DU	-ren	n-Ø	-Ø	-ben	n- -f	-f	-fori	n- -fi	-fi
PL	-men	n-(rV)m	-(rV)m	-reben	n- -(rV)f	(rV)f	-rofori	n- -(rV)fi	-(rV)fi

(11) Class 2 Actor affixes in Maori:

	REAL/MacroPRES			NrPST (Imperfective)			RmPST (Imperfective)		
	1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
SG	-du	-Ø	-Ø	-men	-m	-m	-maf	-maf	-maf
DU	-den	n-Ø	-Ø	-men	n- -m	-m	-maf	n- -maf	-maf
PL	-men	n-Ø	-Ø	-ben	n- -b/-m	-b/-m	-baf	n- -baf	-baf

Crosscutting number and person categories are two aspect features, completive/perfective (class 1) vs. imperfective (class 2), and different tenses (remote past, near/non-remote past and present/future). The aspectual property of the stem is impor-

---

4 The plural *-rVm* appears only with intransitive subjects (S), typically in middle voice.

tant for the selection of the relevant suffix. Thus, verbs expressing punctual/telic events such as 'fall (onto X)' take the completive suffix *-f*, rather than the imperfective *-m* for the non-remote past: *sorono-f/*sorono-m* 'fall-3NPL.NrPST'. In contrast, non-punctual durative events like 'stay' take class 2 NrPST *-m* rather than *-f*, producing *kuye-m*, 'stay-2/3.NPL.NrPST'; the form **kuye-f* does not exist.

It should be noted that all forms of the A suffixes encoding a three-way distinction are in fact constructed from possibly underspecified exponents. For example, the first person future Class 1 suffixes *-ru/-ren/-men* and Class 2 suffixes *-du/-den/-men* consist of the following formatives:

(12) Class 1 and Class 2 first person future suffixes:

r/d	'NPL'	vs.	*m*	'PL'
u	'SG'	vs.	*e*	'NSG'
Ø	'SG'	vs.	*n*	'NSG'

Likewise, the second person future affixes make use of discontinuous *n-* 'NSG' and *-(rV)m* 'PL' (in contrast to *-Ø* 'NPL').

Class 1 A suffixes also include the plural suffix *-rV* with the allomorph *-fre*.[5] This suffix is used to encode plurality of participants and/or plurality of states/inchoative events (typically inchoative states such as 'be(come) dead') and telic/punctual events such as 'fall', 'leave (for)', 'kick'). Thus, the auxiliary for the sentence 'they were dead' is inflected with the plural A suffix *-rV* realized as *-fro* in *nggo(r)-fro-f* 'AUX-PL-3NrPST'.

Verbs in imperatives in Marori are inflected with number indexing following the IRR/FUT form (see Table (10)), and marked by the particle *ka*. In the following examples, the verb 'sit' shows verbal number:

(13) a. ka=mi!
      2.IMP=SG.sit
      'You (SG) sit!'

   b. ka=nermi!
      2.IMP=DU.sit
      'You two sit!'

   c. ka=minenggem!
      2.IMP=PL.sit
      'You (PL) sit!'

---

5 The allomorph *-fre* appears in the morphophonological context of the past suffix *-f(i)*, which follows a heavy syllable. For example, in *ngge+rV+f* 'AUX-PLvb-NrPST' , *-rV* becomes *-fre* (i.e., *nggefref*) because the stress falls on the first syllable, making it a heavy syllable needing a coda which is filled in by the available consonant material *-f*.

1.4 Role flagging and number

The basic clause structure of Marori is shown in (14). The verb in Marori is typically clause-final. Free NP dependants are not obligatory (as indicated by the star *) and have no fixed order; they can also be postposed after the verb in certain cases for information-structural reasons (Arka, 2017). A predicative event noun, if any, precedes the finite verb.

(14) NP* (Event.Noun) Inflected.Verb

Like pronouns, free NPs receive role flagging. The semantic role (theme, patient or recipient) is typically flagged with the invariant clitic =*i*, although the clitic flagging can be dropped. The verb is also marked with a pronominal affix indexing the U role: theme/patient for transitives, recipient for ditransitives. In (15), for example, the patient U argument *na* '1SG' is marked by =*i* and cross-referenced with the prefix *i-*.

(15) Pa=na=i   Thomas ter=i-mo-Ø
     soon=1SG=U Thomas hit=1SG.U-AUX-3NPL.A.FUT
     'Thomas will hit me soon.'

The verb *njime-* 'give' is the only true ditransitive predicate (i.e. with two objects): the recipient bears the U role, indexed by the vowel on the verb. The stem for the verb shows no verbal number. The enclitic =*i* typically marks the recipient U as in (16-a). Both objects can be marked with =*i* as in (16-b-c), however. Crucially, as shown in (16-c), it is only the recipient *Maria* that can be indexed on the verb.

(16) a. Na Albert=i njime-ben   bosik sokodu.
        1SG Albert=U 3M.SG.U.give-1NPL.A.NrPST.PFV pig   one
        'I gave Albert a pig.'

     b. Pafe sorweri=i John njim-im   poyo=i   fis.
        DEF basket=U John 3M.SG.U.give-3A.NrPST.IPFV coconut=U yesterday
        'John filled the basket (with) coconuts yesterday.'

     c. Na fis   njomo-bon   Maria=i bosik=i sokodu.
        1SG yesterday 3F.SG.U.give-1NPL.A.NrPST.PFV Maria=U pig=U   one
        'I already gave Maria a pig yesterday.'

Recipients can also be marked by the invariant postpositional clitic =*na*, which cannot be dropped. Arguably, the *na*-marked recipient is grammatically oblique, as it is not indexed on the verb. As shown in (17), for instance, the 3F.SG verbal prefix indexes the theme 'female pig', not the male recipient *John*. Oblique recipients do not control verbal number marking.

(17) Pa=na   efi yembutel bosik ndo-ru   John=na
     soon=1SG that female   pig   3F.SG.U.carry.SG-1SG.A.FUT John=DAT
     'I will carry a female pig for John.'

Other peripheral roles are flagged by postpostions, e.g. instrument by *ngge*, (insepa-rable-)comitative by *=fa*, and locative with *=ku*, as shown in (18). Roles flagged by postpositions are not indexed on the verb, and the postpositional clitics cannot be dropped.

(18) a. sepeda=fa suo-f   sekola=ku
        bike=COM go-3NPL.A.NrPST school=LOC
        'S/he went to school by bike.'

     b. esa   yoropur   samagau ngge ter=me-f
        village grandfather club   with hit=3M.SG.U.AUX-3NPL.A.NrPST
        'The village chief hit (it/him) with a club.'

     c. Na fis   ti-no=nggo-bon   mbaren   kwi
        1SG yesterday hide-MID=AUX-1NPL.A.NrPST.PFV back.side tree
        moro=ku.
        head=LOC
        'I hid myself behind the tree yesterday.'

2 Pronominal, nominal, and verbal number

2.1 Generalities

As discussed in Section 1.1, the three-way SG/DU/PL distinction in the Marori nomi-nal number system is explicitly marked only for first and second person bound pro-nominals on the verb, and only for the A (actor) argument. Derived predicative ad-jectival nominals, determiners, and some verbal affixes distinguish SG from NSG number, while some parts of the verbal and pronominal paradigms distinguish NPL from PL number. A few nouns are lexically specified for SG or NSG number, but most common nouns are not marked for number.

     Verbal number in Marori is expressed via suppletive roots, by the pluractional suffix *-ro*, or by a combination of these, as we discuss in Section 2.4.

2.2 Pronominal number

Free personal pronouns in Marori show a singular-nonsingular distinction, as shown in Table (19). The same forms are used for different grammatical roles, and

are flagged accordingly by case clitics. For example, the pronoun *na* '1SG' is goal (dative) in (20-a) and locative in (20-b), and is flagged with *=n* and *=ku* respectively.

(19) Free pronouns in Marori:

	1	2	3
	1	2	3
SG	na	ka	efi
NSG	nie	kie	emnde

(20) a. Ka   na=n   manem nambana!
        2IMP 1SG=DAT story   3SG.U.tell.IMP
        'Tell (your) story to me.'

   b. Efi nam-on   kera   na=ku   rapnenje-f.
      3SG POSS-PST disease 1SG=LOC jump-3NPL.A.NrPST
      'His disease infected me (lit. his disease jumped onto me).'

The reflexive pronouns, given in (21), are morphologically complex, composed of three formatives: the pronominal exponent, possessive *-nam*, and intensifier *-ndu*. Except for the third person forms, the pronominal exponents are clearly related to the free pronominal forms in (19). Like their nonreflexive counterparts, they show a singular-nonsingular number distinction. Note that there is a syncretism between the first person nonsingular and the third person singular reflexive forms.

(21) Reflexive pronouns in Marori:

	1	2	3
	1	2	3
SG	nam-ndu	ka-nam-ndu	ni-nam-ndu
NSG	ni-nam-ndu	ki-nam-ndu	a-nam-ndu

Determiners expressing spatial deixis in Marori have SG/NSG forms, in addition to general number (GEN) forms cross-cutting the proximal, semi-distal and distal distinction, as shown in example (22).

(22) Demonstratives and spatial deictics:

	NEAR		AWAY FROM SPKR/ADDRS.	
	SPKR	ADDRS.	SEMI-DISTAL	DISTAL
SG:	kefi	pafi	nggafi	nggwofi
NSG:	kemnde	pamnde	nggamde	nggwomde
GEN:	keme	pame	nggame	nggwome

These determiners can appear alone or with nouns. Note that the proximal forms, *kefi* and *kemnde*, are clearly related to the third person pronouns *efi* and *emnde*

respectively; see Table (19). Indeed, *efi* and *emnde* are also used as determiners: the examples in (23) show *(k)efi* as a determiner (23-a) and a pronoun (23-b). The examples in (24) provide evidence that *keme* can be used in singular (24-a) and plural (24-b) contexts. In (24) we have distinct verb forms for 'carry': *nde* (with a singular object) vs. *kei* (with a non-singular object).

(23) a. tok=efi   botol reruwo rowae kuya-maf.
        frog=3SG.PROX bottle jar   inside BE.2/3NPL.A-2/3.PST
        'The frog stayed inside the jar.'

     b. Kefi   nam   ai   te
        3SG.PROX 1POSS aunt BE.3NPL.PRES
        'This is my aunt.'

(24) a. Keme   ka   nde!
        3.PROX 2IMP 3M.SG.U.carry
        'You carry this!'

     b. Mbia nandu   kei-n   keme=i.
        PROG 1SG.DAT 3NSG.U.carry-DEIC 3.PROX=U
        'He is carrying these only for me.'

2.3 Nominal number

Most common nouns in Marori show general number (i.e., are unspecified for number). Their singular or plural reference is understood from the indexing on the verb. The common noun *awo* 'kangaroo', for example, has singular or nonsingular interpretation depending on the shape of the verb root and the vowel quality of the verb. In (25-a), for example, *awo* is understood as SG as it is co-indexed by the U prefix (a zero prefix) and the high vowel /e/ inducing vowel spreading across the verb. Example (25-b) is actually ambiguous out of context, either F.SG or NSG, because the vowel /o/, which also induces vowel spreading across the word, encodes 3F.SG or 3NSG. We do not explicitly gloss general number for common nouns.

(25) a. Na awo=i   Ø-ife-ben.
        1SG kangaroo.GEN=U 3–3M.SG.U.see-1NPL.A.NrPST.PFV
        'I saw a/the male kangaroo.'

     b. Na awo=I   Ø-yofo-bon.
        1SG kangaroo.GEN=U 3–3F.SG/3NSG.U.see-1NPL.A.NrPST.PFV
        'I saw (the) female kangaroo/kangaroos.'

A few nouns are lexically specified for SG or NSG number, and require SG or NSG indexing accordingly. Such nouns (e.g. *parapur* 'girl.SG'/*moipur* 'boy.SG', *mening-gon* 'children.NSG', see Section 1) tend to refer to humans.

Derived nominals are inflected, showing a SG/NSG distinction, marked by *-en* (or *-on* depending on the vowel of the stem) and *-(n)de* respectively. For example, from the adjective base *kutow* 'bad' we can derive *kutow-on* 'badNMLZ.SG' ('bad one') and *kutow-nde* 'bad-NMLZ.NSG' ('bad ones'). From the root *abon* 'steal', we can derive *abon-on* 'steal-NMLZ.SG' ('thief') and *abon-de* 'steal-NMLZ.NSG' ('thieves'). Examples:

(26) a. kefi   buku wonnggo-won te.
         PROX.SG book good-NMLZ.SG BE.3NPL.PRES
         'This book is good/a good one.'

   b. kemde   buku wonngo-nde   te-re.
         PROX.NSG book good-NMLZ.NSG BE-3PL.PRES
         'These (more than two) books are good (ones).'

2.4 Verbal number

Participant number can also be expressed by verbal number. That is, verbal stems vary depending on the number of the participants involved.

Verbal number is a category of number related to events, typically reflecting the number of tokens/times an action/state happens (Durie 1986; Corbett 2000, Chapter 8; Veselinova 2008). Event number is a complex phenomenon in Marori, intricately related to nominal number. Plural events can be conceptualised as iterated events involving the same participants (on a given occasion or different occasions), or as distributive events in a given space involving different participants.

Verbal number differs from nominal number as encoded via free pronouns and argument indexing along three dimensions: morphology, the Animacy Hierarchy, and semantics. Morphologically, distinctions in pronominal argument marking are expressed by regular/systematic morphology (prefix, suprasegmental vowel harmony, and suffix) whereas verbal number distinctions are expressed by a combination of irregular/suppletive stem alternations and the presence of *-ro*.

Distinctions in pronominal number clearly follow the Animacy Hierarchy:
- a 3-way distinction in the bound pronouns on the verb for first or second person Actor
- a 2-way distinction of NPL vs. PL on the bound pronouns on the verb for third person Actor
- otherwise a 2-way distinction, e.g. in the U prefix, and free pronouns/demonstratives.

Verbal number does not follow commonly accepted versions of the Animacy Hierarchy, which place first person above second person: when we find a 3-way distinction e.g. in the verb 'come' (cf. (37)), the duactional suppletive stem is only found in the second person, not the first person.[6] The duactional in the first person is expressed compositionally, making use of the Actor suffix. The most common pattern of distinctions in verbal number is a two-way distinction contrasting PL vs. non-PL, which applies to all A/U arguments, regardless of person. For example, the PL root *kswe-* 'hit' can be used to express the meanings 'they hit him' or 'he hit them'.

Semantically, pronominal number is strongly governed by the cardinality of referents: SG, DU, PL for 'one', 'two', 'more than two' entities respectively. In contrast, verbal number has a weaker relation to cardinality. While different from aspect, it is tied to aspect, e.g. plural events and duration. Thus, PL verb/events are typically associated with a higher number of events with a recognised duration (i.e. 'being repetitive'). A quick sequence of three or four events might not be considered as being sufficiently durative, and as such might be expressed via an unmarked/nonpluractional verb.

Verbal number in Marori is expressed through the strategies given in (27), further discussed in the following subsections.

(27)   a. By suppletive roots forming inflected verbs;
       b. By suppletive roots of lexical nominal predicates preceding the auxiliary;
       c. By the pluractional suffix *-ro* forming inflected verbs;
       d. By combinations of (a)–(c)

2.4.1 Suppletive root alternation for verbs

The first strategy, suppletive root alternation for inflected verbs, is relevant for a closed class of verbal roots, exemplified by the intransitive verb 'come' in (28) (see example (39)).[7] This root alternates between *umo-* and *seri-*, depending on whether the events are NPL (28-a) or PL (28-b), which correlates with the number of subject participants. Note that there is no material in the AFF[+1] slot in these examples. Root (or stem) alternations for the verb 'come' are more complex than this, however, as we show in Section 2.4.5 on constructed duactionals.

---

6 For discussion of the relative order of first and second person at the top of the Animacy Hierarchy, see Silverstein (1976), Dixon (1979), Siewierska (2004, 149–151), and references therein.
7 The current lexical database shows that there are at least 40 suppletive roots in Marori. The precise nature of their morphosyntactic and semantic properties is a matter for future research.

(28) a. Ami (Mikael fi)   tamba umo-nof
Ami Michael COM already come.NPL-3A.RPST
'Ami (and Michael) already came (here) [a long time ago].'

b. Emnde usindu tamba   seri-nof
3NSG   all   already come.PL-3A.RPST
'They all came (here) [a long time ago].'

2.4.2 Suppletive root alternation for nominal predicates

The second strategy is also relevant only for a small closed class of nominal roots, and involves suppletion. It is exemplified by the predicate 'hit' in (29)–(30). The suppletive roots precede the inflected verb with the root *mV-*. Morphologically these lexical roots are clitics, attaching to the following word as indicated by the notation =, but they are independent syntactic words. As shown in (29), when the U is third person singular, the root also carries gender information: high vowels (*e* or *i*) for '3M.SG' and low vowels for '3F.SG (and others)'. The vowel quality indicating gender spreads across the lexical predicate through vowel harmony within the phonological word.

(29) a. ter=Ø-me-ben
hit.NPL=3U-3M.SG.U.AUX-1NPL.A.NrPST.PFV
'I hit him (once or twice).'

b. tor=Ø-mo-bon
hit.NPL=3U-3F.SG.U.AUX-1NPL.A.NrPST.PFV
'I hit her (once or twice).'

(30) a. keswe=Ø-mi-men
hit.PL=3U-3M.SG.U.AUX-1SG.A.NrPST.IPFV
'I hit him repeatedly.' or 'I was hitting him.'

b. kaswa=Ø-ma-mon
hit.PL=3U-3F.SG/3NSG.U.AUX-1SG.A.NrPST.IPFV
i. "I hit her repeatedly' or 'I was hitting her.'
ii. 'I hit them (two or more)' or 'I was hitting them (two or more).'

An important point to note from examples in (29)–(30) is the interpretation of event number, which in this case consists in an opposition of nonplural vs. plural events. NPL 'hitting' involves a small number of hits, while its PL counterpart can be translated as 'hit repeatedly'. Native speakers would accept three tokens of hitting as non-durative, licensing the use of the NPL root. This is a crucial difference between event number and nominal number, where NPL for nominal number is either SG or

DU. That is, the aspectual nature of the event – in this case the punctual nature of 'hitting' – is critical for the interpretation of event number. In short, a plural event of hitting by a single actor is necessarily repetitive, involving at least several tokens of hitting (typically many more; therefore, it is unsurprising that it also marks progressive aspect, as we discuss in Section 2.4.6).

Note the same root form, realised as *kaswa-* in (30-b), is used to express the plurality of the U participant (participant number, reading (ii) in (30-b)) and a pluractional reading with a feminine U argument, reading (i) in (30-b)). This example highlights syncretism in the paradigm, a common feature in this language.

Finally, different shapes of the roots also trigger the use of different A suffixes, reflecting different aspectual properties of the events, e.g. perfective *-ben* (29-a) and imperfective *-men* aspect (30-a).

2.4.3 Pluractional suffix *-ro*

In the third strategy, event plurality is marked by the pluractional suffix *-ro* (with its allomorphs *-re/-ri/-ra* depending on gender, tense/aspect, and vowel harmony). This is a productive suffix to form plural events:

(31) a. if the stem denotes an event that is telic (e.g. 'nod' as in (32)), then *-ro* signifies a temporally plural event.
   b. if the stem denotes an event that is atelic (e.g. 'be short' as in (33) or 'carry' as in (35)), *-ro* signifies a distributive plural event (with plural participants).

This suffix may look like a subject agreement marker, but it is not. Evidence for this comes from the fact that it can appear with a SG subject, as shown in (32-a). In this example, pluractional *-ri* expresses an iterative meaning, in contrast to (32-b) where it is absent.

(32) a. ke   na kaŋgari umo-ndu,   mbe Thomas fek
       When 1SG here   come-1SG.A exist Thomas nod
       nggu-ri-m.
       AUX-PL-2/3NrPST.IPFV
       'When I came here, Thomas was nodding.'

   b. ke   na kaŋgari umo-ndu,   Thomas fek nggu-f.
       when 1SG here   come-1SG.A Thomas nod AUX-3NrPST
       'When I came here, Thomas nodded (once).'

In the expression of states, the pluractional suffix is necessarily associated with PL participants. This is an instance of a distributive PL state, whereby the subject is pluralised by the pluractional suffix.

(33) a. John sor-on   te.
        John short-SG.NMLZ BE.3NPL.PRES
        'John is short (Lit. John is a short one.).'

   b. emnde usindu sor-de   te-re.
      3NSG   all   short-NSG.NMLZ BE-3PL.PRES
      'They are all short (Lit. they are short ones).'

2.4.4 Pluractional suffix *-ro* and root suppletion

The fourth strategy to express plural events involves the use of the pluractional suffix in combination with the suppletive root. This gives rise to a space in which complex verbal number meaning can be found. The verb 'carry', for example, has the four stems shown in (34) before nominal number morphology is added. For simplicity, only one meaning, namely the distributive plural event with a SG object (cell (iii) in (34)), is exemplified in (35). While SG in form, the U participant ('coconut') is understood as PL, due to the plurality of the event of carrying.

(34) Stem forms of the verb 'carry':

	SG.U	NSG.U
NPL.A	(i) ndV-Ø	(ii) kei-Ø
PL.A	(iii) ndV-rV	(iv) kei-rV

(35) nie   usindu sajer-sajer sokodu poyo=i   nde-re-men
     1NSG   all   day-REDUP   one   coconut=U 3M.SG.U.carry-PL-1PL.PRES
     pambe.
     there
     'We all (three or more), each of us, every day carry one coconut there.'

In (36), verbal number is marked by a combination of root suppletion and the pluractional suffix *ro*: the lexical predicates are expressed by the suppletive roots *anep* 'big.SG' and *kofe* 'big.NSG', and the pluractional suffix appears in (36-b).[8] Note that the inflected auxiliaries come with both U and A affixes, signalling the middle voice in Marori; see Arka (2015) for details.

(36) a. tamba=na   anep   i-nggo-bon.
        already=1SG big.SG 1SG.U-AUX-1NPL.A.NrPST.PFV
        'I have become big.'

---

8 At this stage of our knowledge of Marori, *anep* 'big.SG', *kofe* 'big.NSG', *monjun* 'small/little.SG' and *mnindum* 'small/little.NSG' are the only suppletive property words (categorially nouns) which can appear in predicative/attributive positions without derivation, and impose number agreement.

b. Nie   usindu tamba kofe   yor-nggo-ro-bon.
    1NSG all   already big.NSG 1NSG.U-AUX-PL-1NPL.A.NrPST.PFV
    'We all have become big.'

2.4.5 Constructed duactionals

Duactional number may be encoded by a specific suppletive stem, or constructed by combining NSG and NPL forms. For example, the verb meaning 'come' may have (partially) suppletive stems referring to two tokens of 'coming'. This is illustrated by the paradigm in (37), where the form *nworimo-* is the suppletive duactional form. The pluractional form is *nojri-*. The suffix *-n* is a deictic marker (toward the speaker), also functioning as a past perfective marker. Note that this paradigm is structured according to the Animacy Hierarchy, with a greater number of distinctions higher in the hierarchy.[9]

(37)   Paradigm for the verb 'come':

	FUT/IRR	MacroPRES	RPST
1SG	umo-ndu	umo-ndu	umo-nofori
1DU	umo-nden	umo-nden	umo-nofori
1PL	seri-ndu	seri-ndu	seri-nofori
2SG	umo-n	umo-n	umo-nof
2DU	nworimo-n	nworimo-n	nworimo-nof
2PL	nojri-n	nojri-n	nojri-nof
3NPL	uma-m	umo-n	umo-nof
3PL	ya-m	seri-n	seri-nof

Example (38) exemplifies constructed duactional verbal number. The remote past paradigm of the copula 'be' for the first person is irregular: *orowe* 'BE.1NPL.RPST' vs. *mirrarnggi* 'BE.1PL.RPST'. The root for 'short' is *sor-*, and DU states with DU participants are constructed without specific DU morphology, by a combination of NSG and NPL:

(38) Nie   mundo   sor-de   orowe.
     1NSG long.time.ago short-NSG.NMLZ BE.1NPL.RPST
     'We two were short a long time ago.'

---

9 In (37), MacroPRES covers present (PRES) and non-remote past (NrPST).

2.4.6 Verbal number and aspect

As briefly discussed in Section 2.4.3, verbal number and (grammatical/lexical) aspect are distinct categories in Marori. They are encoded differently and have their own distinct functions in the grammar. Structurally, suppletion for verbal number can be found with roots of inflected forms or with the event noun component of a nominal + light verb combination, whereas grammatical aspect is only expressed by verbal affixation. Functionally, verbal number expresses the number of event tokens and/or number of participants.

Verbal number and aspect are intimately related, however. Forms expressing verbal number show either a SG/NSG distinction or a PL/NPL distinction. The distinction appears to be related to the lexical aspect of the root, i.e. whether the depicted event is aspectually stative or dynamic, or whether it is punctual. Consider a sample of lexical predicates from different lexical classes showing root alternations in (39). The general patterns are as follows: 1-place state predicates (39-a,b) and those expressing non-punctual events (39-f) typically have forms showing a SG/NSG distinction, whereas predicates expressing motion (39-c,d) and punctual action (39-e) show a PL/NPL distinction, keeping in mind the caveat in Section 2.4.2 about a small number of events (possibly more than two) potentially treated as NPL.

(39) a. *anep* 'big.SG' vs. *kofe* 'big.NSG'   (state: 1-place pred.)
      b. *monjun* 'small.SG' vs. *menindum* 'small.NSG' (state: 1-place pred.)
      c. *kunonjo* 'go.NPL' vs. *kurfenj* 'go.PL'   (motion, −telic: 1-place pred.)
      d. *umo* 'come.NPL' vs. *seri* 'come.PL'   (motion, +telic: 1-place pred.)
      e. *trV* 'hit.NPL.U' vs. *kswV* 'hit.PL.U'   (+punctual: two-place pred.)
      f. *ndV* 'carry.SG.U' vs. *kei* 'carry.NSG.U'   (−punctual: two-place pred.)

Certain plural events must appear with (auxiliary) verbs in the imperfective (IPFV) aspect; cf. the contrast in examples (29)–(30) for 'hitting'. Given that 'hitting' is aspectually punctual, this constraint is unsurprising, since such temporally plural events (by the same participants) are necessarily repetitive. However, with other predicates, the reverse does not hold; imperfective aspect does not always require PL repetitive events. Predicates expressing atelic/non-punctual events, such as 'carry', can appear with their NPL root *ndV-*, which co-occurs with imperfective aspect – in this case *-m*, as shown in (40).

(40) Piter Albert=i ndi-m.
      Peter Albert=U 3M.SG.U.carry-3NPL.A.NrPST.IPFV
      'Peter carried Albert.'

In addition, plural events do not necessarily appear in imperfective aspect. Thus, spatial plurality of events (i.e. distributive simultaneous plural events) can appear in perfective aspect, as shown in example (41).

(41) Emnde poyo=i   nde-fre-f   nggambe.
     3NSG   coconut=U 3M.SG.U.carry-PL-NrPST there
  (i) 'They each carried a coconut/coconuts there.' (total: more than three
      coconuts)
  (ii) 'They carried a coconut there.' (i.e. a total of one coconut)

To conclude, examples like (40) (a SG event with a SG participant, imperfective
aspect) and (41) (plural events in perfective aspect) provide strong evidence that
event/verbal number is distinct from aspect.

2.5 Constructed dual number marking

Dual number can be expressed with or without dedicated dual morphology in Marori. Dedicated dual morphology for nominal number is observed only for the first-
person category; see (10)–(11). DU without dual morphology for nominal number is
observed in the second person category. Consider the expression of number in the
future tense (42), exemplified in (43), where two underspecified formatives are used.
DU is constructed from the combination of NSG (*n-*) and NPL (*-Ø*). Note that the
stem *kswV* 'hit.PL' exhibits plural verbal number: 'hit repeatedly'.

(42) Dual number marking for second person, future tense:
                EXPONENT1-ROOT-EXPONENT2
      SINGULAR Ø-   -Ø
      DUAL   n-   -Ø
      PLURAL   n-   -m

(43) a. kesweme
        kswV=Ø-Ø-me-Ø
        hit.PL=3U-2SG.A-3M.SG.U.AUX-2NPL.A
        'You (SG) will hit him (repeatedly).'

     b. kesneme
        kswV=Ø-n-me-Ø
        hit.PL=3U-2NSG.A-3M.SG.U.AUX-2NPL.A
        'You (DU) will hit him (repeatedly).'

     c. kesnemem
        kswV=Ø-n-me-m
        hit.PL=3U-2NSG.A-3M.SG.U.AUX-2PL.A
        'You (PL) will hit him (repeatedly).'

2.6 Number and the comitative/inclusory/associative construction

Bound pronominals on the verb (i.e. affixes occupying positions AFF^{-1} and AFF^{+2} in (3)) are referential.[10] That is, they can independently refer to entities in the discourse without the presence of free argument NPs. The presence of free argument NPs is therefore pragmatically motivated, often by the need for the argument to be focused or topicalised; see Arka (2017) for a detailed discussion of information structure in Marori. The relation between the overt free NP and the bound pronoun on the verb is therefore anaphoric: the free NP typically functions as a TOPIC NP, with an anaphoric relation to the bound pronoun.

Importantly, co-occurrence of units with incompatible values for the NUM feature is permitted in Marori, and with nonsingular verbs this gives rise to a comitative-inclusory interpretation, as discussed in detail by Arka (2019).

In (44), the 3NPL actor suffix *-f* is compatible with the free NP *John*, and the singular reading (i) is found. In contrast, the verb in (45) carries the pluractional suffix (realised as its allomorph *-fre*), and the comitative inclusory plural reading (iii) is found. Notably, an inclusory-associative DU reading (ii) is not possible in (44)–(45), with a singular undergoer as indicated by the vowel *e* in the auxiliary form. To express the inclusory DU reading, the subject argument *John* must be flagged with the comitative marker *fi* and the verb must appear with inclusory NPL actor morphology, as shown in (46). This again highlights the point about distributed exponence of number, allowing DU to be constructed without using specific DU morphology. Note that NPL *-f* is inclusory in (46) in the sense that it refers to a group of two including John, in contrast to (44) where it only indexes the NP 'John' with a singular interpretation.

(44) John kier=i   ki=ngge-f.
     John village=U leave=3M.SG.U.AUX-3NPL.A.NrPST
     (i)   'John left the village.'
     (ii) *'John and his associate left the village.
     (iii) *'John and his associates left the village.'

---

10 The third person A/subject suffix can be non-referential when used with a predicate expressing weather, e.g.:

(i)   tanamba mag   te.
      now   hot   BE.3NPLPRES
      'It is hot now.'

(45) John kier=i   ki=ngge-fre-fi.
     John village=U leave=3M.SG.U.AUX-PL-3A.RPST
     (i)   *'John left the village.'
     (ii)   *'John and his associate left the village.'
     (iii) 'John and his associates left the village.'

(46) John fi   kier=i   ki=ngge-f.
     John COM village=U leave=3M.SG.U.AUX-3NPL.A.NrPST
     (i)   *'John left the village.'
     (ii)   'With John (included), they (two) left the village.'
     (iii) *'With John, they left the village.'

Free pronouns can also function inclusorily. In (47), *nie* '1NSG' is the first person NSG pronoun, indexed on the verb with the actor suffix *-den* '1DU.PRES'. As seen from the translation, the combination *nie...-den* refers to a totality of two participants in the event of walking, including the speaker, and cannot refer to three or more people.

(47) nie   [bab desa   fi]   keme uma-den   mukedu.
     1NSG uncle village COM REL   walk-1DU.A.PRES middle
     '(Here are) the village chief and I who walk in the middle.'

Proper names can appear without the comitative marker *fi*, as in (45). Free pronoun arguments require *fi*, as shown in (48).

(48) Efi   *(fi) kier=i   ki=ngge-fre-fi.
     3SG COM village=U leave=3M.SG.U.AUX-PL-3A.RPST
     'With him/her, they left the village.' (at least three people left)

The inclusory-comitative construction with mismatched number can also involve inanimate nouns, which are not marked for number. For example, the plural verb *kei* 'carry' requires a nonsingular U argument, but it can appear with a singular NP with singularity encoded by the modifier *anep* 'big.SG' in (49). As seen from the translation, the understood U participants are plural, not singular, with the 'single big coconut' being included in the plural set.

(49) Fis   anep poyo=i   sokodu Ø-kei-ben.
     Yesterday big.SG coconut=U one   3U-carry.NSG.U-1NPL.NrPST.PFV
     'One big coconut, I carried it yesterday with the other ones/things.'

The reverse indexing pattern with an animate plural free NP and a singular verb does not give rise to an inclusory meaning. Rather, it encodes a 'small plural' or

paucal-like meaning. Consider (50), where the pluralia tantum noun *meninggon* 'children' is indexed with a nonplural Actor suffix *-f*, giving rise to the meaning 'a small number of children'.

(50) Meninggon kier=i   ki=ngge-f.
    children   village=U leave=3M.SG.U.AUX-3NPL.A.RPST
    'A small number of children (i.e. a few, two or three) left the village.'

3 Agreement and the syntax of number

3.1 Argument indexing and role flagging

Marori does not exhibit verb agreement with nominal arguments; rather, bound pronominal arguments are indexed on the verb, and characterized by complex patterns of distributed underspecified exponence. Argument indexing and role flagging are discussed in Sections 1.3 and 1.4. Mismatches in number morphology can also indicate particular semantic distinctions: for example, a singular NP indexed by a plural verb expresses an inclusory meaning, and a plural NP indexed by a singular verb expressed a 'small plural'/paucal meaning, as discussed in Section 2.6.

3.2 Number within NP

The Marori nominal phrase is a determiner phrase (DP: Abney 1987), as shown in (51-a); a determiner (D) itself can stand alone as a nominal phrase, or can take a NP complement. Furthermore, the determiner is optional within the DP, and so a determinerless NP can also stand on its own as a nominal phrase. The determiner, one of the spatial deictics shown in (22), and the noun phrase (NP) within DP can be freely ordered, as indicated by the comma in (51-a). The NP itself, as shown in (51-b), can consist of a possessive phrase (PossP), numeral/quantifier (Num/Quan), adjective (nominal) (Adj(N)) and a noun. These elements are freely ordered. It follows from the DP structures in (51) that a determiner cannot appear within the NP; it must appear initially or finally in the nominal structure. This is exemplified in (52), in which (52-c) is ungrammatical because the determiner *nggafi* is inside the NP.

(51) a. [NP, D]$_{DP}$

    b. [PossP, Num/Quan, Adj(N), N]$_{NP}$

(52) a. [Nggafi [Thomas nam sour]$_{NP}$]$_{DP}$
        DIST.SG Thomas POSS house

b. [[Thomas nam sour]$_{NP}$ nggafi]$_{DP}$
   Thomas   POSS house   DIST.SG

c. *[[Thomas nam nggafi   sour]$_{NP}$]$_{DP}$
   Thomas   POSS DIST.SG house
   'That house of Thomas'

Number within the nominal domain is relevant for the following categories: determiners (e.g. *kefi* 'SG' vs. *kemnde* 'NSG', see (22)), certain property words (e.g. *anep* 'big.SG' vs. *kofe* 'big.NSG'), derived adjectival nominals (*-on* 'SG.NMLZ' vs. *-de* 'NSG.NMLZ'), and certain nouns (e.g. *moipur* 'boy.SG'/*yarapur* 'girl.SG' vs. *meninggon* 'children.NSG'). Number agreement is required between these elements if they co-occur, particularly between the Determiner (i.e., head of the nominal phrase) and the head noun of the NP, or between the noun and its modifier, if the modifier happens to be a derived nominal inflected for number. Numerals/quantifiers, because of their number-related meanings, also impose co-occurrence restrictions. In many cases, since the majority of nouns in Marori are not inflected for number, number agreement within NP is not evident. For example, the demonstrative with general number, *keme*, can co-occur with lexically singular or plural nouns: *keme yarapur* 'this girl', *keme meninggon* 'these children'. In what follows, we discuss examples involving nouns that are lexically singular or plural and where number is also overtly marked elsewhere in the noun phrase.

The examples in (53) show determiner-noun number agreement; e.g. the SG determiner *kefi* co-occurs with the singular noun *moipur* 'boy' (53-a). Likewise, the NSG determiner must agree with a NSG noun (53-b). The adjective modifiers *kofe* and *anep* must also appear with the correct number agreement, as shown in (54).

(53) a. moipur kefi/*kemnde
       boy.SG PROX.SG/NSG
       'the/this boy'

   b. meninggon kemde/*kefi
      child.NSG   PROX.NSG/SG
      'the/these children'

(54) a. kemde   kofe   purfur
       PROX.NSG big.NSG child.NSG
       'these adults/elders'

   b. *kefi   kofe   purfur
       PROX.SG big.NSG child.NSG

   c. *kemde   anep   purfur
       PROX.NSG big.SG child.NSG

Example (55) shows that the presence of a numeral also requires number compatibility with the relevant units: *yanadu* 'two' must co-occur with the NSG determiner *pamnde*, not SG *pafi*.

(55) yanadu ujif pamnde   / *pafi
    two   bird DIST.NSG   DIST.SG
    'those two birds'

The examples in (56) involve the adjectival nominalisers *-en* 'SG.NMLZ'/*-nde* 'NSG.NMLZ'. In these examples, the whole modifier units float away postverbally. Since they are associated with U NPs, the SG/NSG NPs require SG/NSG U indexing on the verb.

(56) a. na [sokodu purfam=i]ₙₚ ife-ben   [kaeri
       1SG one   person=U   3M.SG.U.see-1NPL.NrPST.PFV long
       pu-en]
       hair-SG.NMLZ
       'I saw one long-haired person.'

    b. na [yanadu purfam=i]ₙₚ yofo-bon   [kaeri
       1SG two   person=U   3M.NSG.U.see-1NPL.NrPST.PFV long
       pu-nde]
       hair-NSG.NMLZ
       'I saw two long-haired persons'

4 Semantics and Discourse

4.1 Kinds of plurality and the count/mass distinction

Nominal and verbal number in Marori exhibit an intricate interaction with the count/mass distinction. Mass nouns, e.g. *mim* 'water', are like other nouns in having general number, allowing SG or NSG reference. They are treated as (masculine) SG in Marori by default, and accordingly receive SG indexing. However, verbal PL indexing is possible, giving rise to an individuated interpretation, e.g. with the additional meaning 'kinds of', 'pieces of', 'bottles of', etc. as shown in (57). In (57-a), the mass noun *mim* 'water' is the object of the PL verb *kei-* 'NSG.U.carry'. As seen from the free translation, *mim* is understood as referring to individuated quantities, e.g. three or more containers of water. Likewise, the oil in (57-b) is understood as PL (i.e. three or more bottles of oil).

(57) a. Emnde usindu mim=i   kei-fref.
　　　 3NSG   all   water=U 3NSG.U.carry-2/3PL.A.NrPAST.PERF
　　　 'They all carried water (i.e. all carried their own bottles).'

　　 b. poyo   holoi menindun te-re.
　　　 coconut oil   small.NSG BE-2/3PL.PRES
　　　 'They're coconut oil in small quantities (e.g. in (small) bottles).'

The nominaliser *-en* 'SG.NMLZ' and *-(n)de* 'NSG.NMLZ' turns a (mass nominal) stem to a countable nominal. Thus, from the non-count colour noun *singgu* 'black(ness)', we can derive a countable noun *singgu-en* 'the black one' and *singgu-nde* 'the black ones' (two or more), and from the mass event noun *abon* 'steal', we can derive *abon-on* 'the thief' and *abon-de* 'thieves'. Once individuated, these nominals behave as standard countable nouns. This is exemplified in (58), where number is specified through a combination of nominal and verbal number mechanisms.

(58) a. singgu-en   te.
　　　 black-SG.NMLZ BE.3NPL.PRES
　　　 'It's black (or a black one).'

　　 b. singgu-nde   te.
　　　 black-NSG.NMLZ BE.3NPL.PRES
　　　 'They (two) are black (ones).'

　　 c. singgu-nde   tere.
　　　 black-NSG.NMLZ BE.3PL.PRES
　　　 'They (three or more) are black (ones).'

Plural events can be spatially distributive or non-distributive. The same predicate can have both interpretations. This is exemplified by the predicate 'carry', which is lexically atelic, and pluractional affixation (realised by *-fre*) encodes the individuation/pluralisation of the event with plural actors. As shown in (59), the interpretation with the SG object is ambiguous between a distributive plural meaning (i.e. a plural event, reading (i)), and a non-distributive meaning (a single event with plural agents, reading (ii)).

(59) Emnde poyo=i   nde-fref   nggambe (= (41))
　　 3NSG   coconut=U 3M.SG.U.carry-2/3PL.A.NrPST.PERF there
　　 (i) 'They each carried a coconut/coconuts there.' (total: more than three co-
　　　　 conuts)
　　 (ii) 'They carried a coconut there.' (i.e. a total of one coconut)

Note that, with an inherently punctual predicate like 'hit', event plurality can also be interpreted temporally (i.e. repetitive subevents in a single durative occasion by

the same actor; cf. example (30)), in which case, a PL verbal root is used. In (60-a), we see a more complex temporal plurality (also with a PL root) across occasions in durative habitual aspect. Even more complex is the combination of multiple plural events in both spatial and temporal dimensions. This is exemplified by the habitual hitting with PL root as shown in (60-b).

(60) a. Mbe nggewendu John Maria=i   kaswa=Ø-ma-Ø.
       MBE often   John Maria=U hit.PL=3U-3F.U.AUX-3NPL.A.PRES
       'John often hits Maria (with multiple hittings on each occasion).'

   b. Emde usindu mbe Maria=i   nggewendu koswo=Ø-mb-ro.
      3NSG all   MBE Maria=U often   hit.PL=3U-3F.A.AUX-3PL.A.PRES
      'They all often hit Maria (with each of the actors regularly hitting her repeatedly on each occasion).'

4.2 Exclusive and inclusive plural readings

Since the Marori number system exhibits a three-way distinction, the meaning of PL in this language (i.e. 'three or more', or '>2') is not the same as that in English (i.e. 'two or more'). Nevertheless, as in English, Marori number shows exclusive and inclusive plural readings in similar distributional contexts (Farkas & de Swart, 2010). On an exclusive reading, Marori PL forms refer to three or more individuals, while on an inclusive reading, the same forms refer to any number of individuals, including single individuals. Inclusive and exclusive readings of PL noun phrases in English are illustrated in (61-a) and (61-b) respectively. Note that the difference in the interpretation of the PL form 'children' correlates with a difference in clausal types. In the positive context (61-a), the sentence is false if the speaker saw only one child: this is an exclusive plural reading. Contrastively, 'children' in (61-b) has an inclusive plural reading, and its reference includes single individuals.

(61) a. I saw children.
        (Exclusive plural reading only: speaker saw more than one child; false if the speaker saw only one child)
   b. I didn't see children.
      (Inclusive plural reading: speaker did not see any children, not even one).

We find the parallel pattern in Marori, even though the way plurality is expressed in Marori is significantly different from that in English. Consider (62), where the pluractional -*re* encodes the A, i.e 'snake', as PL.

(62) a. John=i kaf   imbi-ref   paya-ke   fis.
        John=U snake 3M.SG.U.bite-2/3PL.A.NrPST.PERF forest-LOC yesterday
        'John was bitten by snakes (> 2) in the forest yesterday.'
        (The exact number of snakes is unknown, but more than two.)

   b. Maar tanamba kaf   John=i imbi-ref
        NEG   just.now snake John=U 3M.SG.U.bite-2/3PL.A.NrPST.PERF
        'No snakes bit John just now.'
        (inclusive: Not even one snake (or two snakes) bit John.)

In a positive context (62-a), the PL receives an exclusive reading involving at least three snakes ('>2'). In a negative context (62-b), the PL has an inclusive reading, as seen from the free translation. In addition to negation, inclusive plural interpretation is also observed in questions and conditionals; see Arka & Dalrymple (2016) for details.

5 Conclusion

The number system of Marori shows distributed marking across different morphological and syntactic sites. The marking regulates the interaction of nominal number (plurality of entities) and verbal number (plurality of events). Nominal number is expressed via marking of free pronouns and derived nouns and the associated verbal indexing. Nominal number shows a three-way distinction and follows the Animacy Hierarchy, with a singular-dual-plural distinction relevant only for first and second bound pronominals (i.e. in the top segments of the hierarchy). Elsewhere, nominal number marking typically shows a two-way underspecified distinction (singular vs. nonsingular, plural vs. nonplural). Verbal number is, in contrast, expressed by suppletive verbal root alternations, not following the animacy hierarchy, and typically showing a two-way weak singular vs. nonsingular distinction.

    An unusual characteristic of the Marori number system is its distributed underspecified exponence in marking, allowing a specific category in nominal/verbal number to be constructed without a dedicated number marker for that category. Thus, dual (nominal number) is expressed not by dual morphology, but by a combination of nonsingular and nonplural exponents; e.g. as seen in the expression of the second person dual (future) *n-…-Ø* in (42). Duactional (verbal) number is likewise expressed in the complex predicate construction by combining a nonplural (auxiliary/light) verb and a nonsingular nominal predicate, as seen in (38).

    The whole system of number underspecification, verbal indexing and verbal number in Marori renders number in this language functionally semantic and pragmatic in nature, quite different from syntactic number in familiar languages like English. For example, a singular dependent NP can be indexed by a plural verb to

express an inclusory meaning, as seen in (45). However, the reverse pattern of co-indexing of a plural free NP and singular verb gives rise to a different meaning, namely 'small plural/paucal' as seen in (50).

In Marori's three-way number system, the meaning of plural is not quite the same as in a two-way number system like English: 'three or more' (Marori) vs. 'two or more' (English). However, while Marori is radically different from English in its number system, it exhibits similarity in terms of the distribution of plural meanings under negation and other contexts (Farkas & de Swart, 2010); inclusive and exclusive plural readings are found in the same distributional contexts as in English and other European languages with a two-way number system.

Acknowledgements

Abbreviations

Glosses follow the Leipzig Glossing Rules (Bickel et al., 2015). We use the following abbreviations:

1	first person
2	second person
3	third person
A	actor
AUX	auxiliary verb
BE	copula
COM	comitative
DAT	dative
DEF	definite
DEIC	deictic/approaching speaker
DIST	distal
DU	dual
F	feminine
FUT	future
GEN	general/unspecified number
IMP	imperative
IPFV	imperfective

LOC	locative
M	masculine
MBE	particle *mbe* indicating start of new discourse episode
MID	middle
NEG	negation
NFUT	nonfuture
NMLZ	nominalization
NPL	nonplural/2 or fewer
NrPST	near/non-remote past
NSG	nonsingular/more than 1
PFV	perfective
PL	plural/3 or more
POSS	possessive
PRES	present
PROG	progressive
PROX	proximal
PST	past
REDUP	reduplication
RPST	remote past
SG	singular
U	undergoer

References

Abney, Stephen P. 1987. *The English noun phrase in its sentential aspect*: Massachusetts Institute of Technology dissertation.

Arka, I Wayan. 2011. Constructive number systems in Marori and beyond. In Miriam Butt & Tracy Holloway King (eds.), *On-line proceedings of the LFG2011 conference*, Stanford: CSLI Publications. http://csli-publications.stanford.edu/LFG/16/lfg11.html.

Arka, I Wayan. 2012a. Projecting morphology and agreement in Marori, an isolate of Southern New Guinea. In Nicholas Evans & Marian Klamer (eds.), *Melanesian languages on the edge of Asia: Challenges for the 21st century* Language Documentation & Conservation, 150–173. Honolulu: University of Hawai'i Press. Special Publication No. 5.

Arka, I Wayan. 2012b. Verbal number, argument number and plural events in Marori. In Miriam Butt & Tracy Holloway King (eds.), *On-line proceedings of the LFG2012 conference*, Stanford: CSLI Publications. http://csli-publications.stanford.edu/LFG/17/lfg12.html.

Arka, I Wayan. 2015. Constructed middles in Marori: An LFG analysis. In Miriam Butt & Tracy Holloway King (eds.), *On-line proceedings of the LFG2015 conference*, Stanford: CSLI Publications. http://csli-publications.stanford.edu/LFG/20/lfg15.html.

Arka, I Wayan. 2017. Information structure and its morphosyntactic resources in Marori. In A. Utsumi & A. Shiohara (eds.), *Proceedings of the third international workshop on information structure of the Austronesian languages*, Tokyo: ILCAA, TUFS.

Arka, I Wayan. 2019. Comitative-inclusory constructions in Marori. In Miriam Butt, Tracy Holloway King & Ida Toivonen (eds.), *On-line proceedings of the LFG2019 conference*, Stanford: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2019/index.shtml.

Arka, I Wayan & Mary Dalrymple. 2016. Number and plural semantics: Evidence from Marori. *NUSA: Linguistic studies of languages in and around Indonesia* 60(30). 89–106. http://www.aa.tufs.ac.jp/en/publications/nusa/back-issue/60.

Bickel, Balthasar, Bernard Comrie & Martin Haspelmath. 2015. Leipzig glossing rules: Conventions for interlinear morpheme-by-morpheme glosses. http://www.eva.mpg.de/lingua/resources/glossing-rules.php.

Corbett, Greville G. 2000. *Number*. Cambridge, UK: Cambridge University Press.

Dixon, Robert M. W. 1979. Ergativity. *Language* 55. 59–138.

Durie, Mark. 1986. The grammaticization of number as a verbal category. In Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj & Deborah Feder (eds.), *Proceedings of the 12th annual meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society.

Evans, Nicholas, Wayan Arka, Matthew Carroll, Yun Jung Choi, Christian Döhler, Volker Gast, Eri Kashima, Emil Mittag, Bruno Olsson, Kyla Quinn, Dineke Schokkin, Philip Tama, Charlotte van Tongeren & Jeff Siegel. 2017. The languages of Southern New Guinea. In Bill Palmer (ed.), *The languages and linguistics of the New Guinea area: A comprehensive guide*, 641–774. Berlin: De Gruyter Mouton.

Farkas, Donka & Henriëtte E. de Swart. 2010. The semantics and pragmatics of plurals. *Semantics and Pragmatics* 3. 1–54.

Haspelmath, Martin. 2019. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. Special issue: Issue in Honour of Frantisek Lichtenberk.

Siewierska, Anna. 2004. *Person*. Cambridge, UK: Cambridge University Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies. Linguistic Series No. 22.

Veselinova, Ljuba N. 2008. Verbal number and suppletion. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures Online*, Munich: Max Planck Digital Library. http://wals.info/.

V Americas

Olga Krasnoukhova

17 Number in the languages of South America

Abstract: This chapter offers a typological overview of the number category in the indigenous languages of South America (SA). The focus is placed on number in independent personal pronouns and on nouns, as well as on verbal number. The discussion is centered around tendencies and patterns that SA languages show in number marking in the (pro)nominal and verbal domains. Whenever relevant, SA data are situated in a larger cross-linguistic context.

The section on verbal number constitutes a first comparative account of this phenomenon in SA languages. It is shown that both types of verbal number (i.e. event number and participant number) are widespread in SA, occurring in most language families surveyed. Teasing apart verbal number (of the participant plurality type) and nominal number manifested on the verb (argument indexing) turns out to be an interesting challenge for SA, with many intermediate cases.

1 Overview

There are about 420 languages in South America (SA). These fall into 53 language families with more than one member and 55 one-member language families, or isolates (Campbell 2012a: 59). Only eight language families are relatively large (with approximate number of members, according to Hammarström et al. 2020): Arawakan (76), Tupían (71), Pano-Tacanan (45), Quechuan (44), Cariban (42), Macro-Ge (30), Chibchan (27) and Tukanoan (26). Most other families are small and comprise less than ten members (Campbell 2012a: 59). SA is also home to about 35 % of the world's isolates (Campbell 2017: 9; Seifart and Hammarström 2017: 260). With this linguistic diversity, structural variation related to number is not surprising. Nevertheless, some general tendencies and patterns can be identified.

The chapter deals with pronominal, nominal, and verbal number. (Pro)nominal number is conceptually associated with the number of participants or entities. 'Pronominal number' refers here to number encoded on free personal pronouns (discussed in Section 2.2). 'Nominal number' is reserved here for encoding of number on the noun or the noun phrase (NP) (Section 2.3). Finally, 'verbal number' is used here as a cover term for encoding of plurality of events or states and participant plurality expressed on the verb (Section 2.4).

Observations on (pro)nominal number are based on a sample of ±100 languages (the total number depends on the feature). Genealogically these fall into about 36 language families with more than two members, and 16 language isolates. Obser-

vations on verbal number are based on a sample of 70 languages (overlapping with the former to a great extent). Genealogically they represent 31 language families with more than two members and 10 isolates.

2 Pronominal, nominal, and verbal number

2.1 Generalities

Asymmetries concerning 'number' encoded in pronominal, nominal, and verbal domain are not rare. First of all, this concerns the sheer existence of the number category in a domain. All SA languages in the sample adhere to the cross-linguistic tendency that if a language has a number category, it will be found in pronouns (cf. Croft [1990] 2003: 129). Within the pronominal domain, the SA languages show that the 1st and 2nd person pronouns are more often marked for number than the 3rd person pronouns. Thus SA data conform to the Person Hierarchy (1st > 2nd > 3rd) and the Nominal Hierarchy (pronouns>nouns), where an item higher (i.e. to the left) on a hierarchy is more likely to be marked than one to the right of it (see Croft 2003: 130; Corbett 2000: 62).

Asymmetries in values across domains (e.g. singular-dual-plural distinction in pronouns, but only singular-plural in nouns) are found as well, and I discuss these in Section 2.5. Verbal number in SA deserves particular attention. Corbett (2000: 245) discusses the geographic extent of verbal number, but, surprisingly, the continent of SA is not mentioned at all. As will be shown, verbal number is a widespread phenomenon in SA, found in many language families.

2.2 Pronominal number

2.2.1 Number distinctions

Pronominal number involves independent personal pronouns. Number distinctions on pronouns are found in a substantial majority of SA languages. This is in contrast with the nominal domain: in a substantial portion of SA languages number on nouns is either optional or impossible.

In the domain of personal pronouns the SA languages show a split: 1st and 2nd person pronouns are more often marked for number than 3rd persons. For example, in Amarakaeri (Van linden, forthc.) and in the Tupían languages Karitiâna (Everett 2006: 303) and Tapieté (González 2005: 121), 3rd person pronoun is number-neutral, whereas 2nd and 1st person distinguish singular vs. non-singular forms.[1]

---

1 Karitiâna and Tapieté additionally have an inclusive-exclusive distinction in 1st person pronouns. This is not the case in Amarakaeri.

In Sabanê, 2nd and 3rd pronouns are number-neutral and only 1st person distinguishes singular vs. non-singular or plural[2] (Araujo 2004: 167). Example (1) illustrates the system in Amarakaeri.

(1) Amarakaeri (Harakmbut; Van linden forthc.)

	Singular	Non-singular
1	**nd?**	**oro?**
2	**on**	**opudn**
3	**ken**	

There are also languages in SA in which all personal pronouns are ambiguous with respect to number. The case of Pirahã (Muran) is an often-cited one: the language has just one form for each person, which can refer to either singular or plural (Everett 1986: 280–283). Plural forms are expressed periphrastically (e.g. 'you and I' for 'we', or 'you and he' for 'you plural') (*idem*.). Besides Pirahã, pronouns unmarked for number are found in Matsés (Panoan) (Fleck 2003: 242–243). Just like in Pirahã, there are means to disambiguate, when necessary. Thus in Matsés the basic pronouns can combine either with the pronoun-specific enclitic *=ben* 'alone/in vain/ for no reason' or with one of the two quantitative postpositions, viz. *daedi* 'both of' or *tedi* 'all of' to produce forms approximating the semantics of dual and plural pronouns. There exists one form with plural semantics, *mitso* '2nd person plural', but it is noted to be archaic (Fleck 2003: 244, 241). Qawasqar is another example of a language where all independent pronouns are ambiguous with respect to number (Clairis 1985: 201; Aguilera 2001: 33; see also Cysouw 2013).[3] One of the means to disambiguate is a pluralizing suffix (-*atal*) on the verb (see Aguilera 2001: 33).

The most common number distinction in pronouns in SA is **singular vs. non-singular**. In 72 % (72 out of 100 languages surveyed – see Appendix 1) singular vs. non-singular distinction is found in all three personal pronouns. In another 7 %, only 1st and 2nd person have singular vs. non-singular, while the 3rd person is

---

2 There is no inclusive-exclusive distinction in Sabanê. Note that the term 'plural', particularly in reference to 1st persons, is semantically awkward (Cysouw 2001: 67, 296). Cysouw (2001), for example, shows that it is more adequate to emphasize not the number but the kind of participants in a group. In this regard it is important to stress that pronominal number is argued to be different from nominal number, as the former lacks 'referential homogeneity' characteristic of the latter (Daniel 2005: 8).
3 Qawasqar is also interesting as the 2nd and 3rd person categories are conflated (cf. Cysouw 2003: 44).

Tab. 1: Languages with plural pronominal forms derived for all 3 persons with a nominal plural marker.

Family	Language	Source
Barbacoan	Tsafiki	Dickinson (2002: 65)
Nambiquaran	Mamaindê	Eberhard (2009: 359)
Matacoan	Nivaclé, Wichí	Fabre (2016: 102); Terraza (2009: 100)
Macro-Ge	Apinajé, Akwẽ-Xerente, Karajá	Oliveira (2005: 159); Filho (2007: 120); Ribeiro (2012: 38)
Quechuan	various Quechuan languages	e.g. Weber (1989:37), see Krasnoukhova, forthc.
Aymaran	Aymara, Jaqaru	Hardman (2001: 101); Coler, forthc.
mixed language	Kallawaya	Muysken, forthc.
isolate	Trumai	Guirardello (1999: 47)
isolate	Kanoê	Bacelar (2004: 141, 118)
isolate	Páez	Jung (2008: 132)

number-neutral.[4] Figure 1 shows the geographic distribution of the number distinctions in the sample (see below).

Quite a few SA languages have non-singular or plural pronominal forms which are morphologically derived for all 3 persons by a nominal plural marker. Cysouw (2003: 70) notes that pronominal paradigms with nominal strategies to mark plurality are cross-linguistically uncommon. For example, in Daniel (2013) only 19 out of 261 sample languages (= 7 %) show this pattern.[5] In SA, we find this strategy in different languages. These are listed in Table 1. Note that in Kanoê, non-singular personal pronouns are formed by a collective marker, which is otherwise found sporadically on some inanimate and animate nouns (Bacelar 2004: 141, 118). In Páez, non-singular pronominal forms are also formed with a collective marker, whose use is limited to human referents forming a group (Jung 2008: 132).

In many other SA languages, in as far as it is possible to generalize, person(s) higher on the person hierarchy often have non-singular forms which are unrelated (or, at least, not straightforwardly related) to singular forms synchronically. See example (1) above. Another example is Yurakaré: 3rd person non-singular is com-

---

4 In one language (Sabanê) only the 1st person distinguishes singular-plural and 2nd and 3rd are number-neutral. Another three languages (Matsés, Qawasqar, and Pirahã) do not show number in pronouns. The remaining 17 languages have singular-dual-plural opposition, as discussed below.
5 The study considers only 1st and 2nd person, inclusives and 3rd person pronouns are excluded from consideration on methodological grounds (see Daniel 2013).

Tab. 2: Languages with dual number in 1st, 2nd and 3rd person independent pronouns.

Family	Language	Source
Araucanian	Mapudungun	Zúñiga (2000: 13)
Boran	Miraña	Seifart (2004: 55)
Chibchan	Chimila, Malayo	Trillos (1996: 121); Trillos (1999: 76)
Chonan	Tehuelche	Fernández Garay and Hernández (2006: 15–16)
Pano-Tacanan	Kashibo-Kakataibo, Cavineña, Araona	Zariquiey (2011: 221); Guillaume (2008: 571); Emkow (2006: 254)
Peba-Yagua	Yagua	Payne and Payne (1990: 370)
Witotoan	Murui	Wojtylak (2017: 154)
Yanomamic	Yanomami, Ninam	Perri Ferreira (2017: 17 5–176); Goodwin Gomez (1990: 50)
Zaparoan	Arabela	Rich (1975: 5–6)
isolate	Camsá	O'Brien (2018: 137)
isolate	Puelche	Casamiquela (1983: 51)
isolate	Trumai	Guirardello (1999: 47)
isolate	Waorani	Peeke (1979: 40)

posed of a singular form and a nominal plural marker, while 1st and 2nd person non-singular are suppletive (van Gijn 2006: 116). A 'split' higher on the person hierarchy is found in Mosetén (Sakel 2004: 117) and Kulina (Dienst 2014: 173), where the non-singular form is suppletive only for 1st person, and the non-singular forms for 2nd and 3rd person are derived with a nominal plural.

We also find languages in which non-singular pronouns are suppletive in all 3 persons: Yuqui (Villafane 2004: 70) and Yanesha' (Duff-Tripp 1997: 61).

Besides the singular vs. non-singular opposition, the second most frequent pattern in SA is **singular-dual-plural** opposition in independent pronouns. We find it in 17 % of SA languages (i.e. 17 out of 100 sample languages). Table 2 presents a list of the languages with the dual distinction in all 3 persons. Figure 1 illustrates the geographic distribution of these languages. As for formal marking, Trumai uses the same dual and plural markers for pronouns as for nouns. In Camsá, this is the case only for 3rd person. In some languages, like Tehuelche, Cavineña, Araona, Yanomami and Ninam, non-singular pronouns contain markers which are formally close to nominal dual and plural markers. In Miraña and Yagua, these are different. The other languages either do not have dual for nouns (Kashibo-Kakataibo, Waorani, Malayo and Chimila) or do not mark any number on nouns at all (Mapudungun and Puelche).

Tab. 3: Occurrence of number distinctions in independent personal pronouns.

Singular – plural distinction in all 3 persons	72%
Singular – dual – plural distinction in all 3 persons	17%
Singular vs. plural in 1st and 2nd person; no number distinction in 3rd person	7%
No number distinction in any person	3%
Singular vs. plural in 1st person; no number distinction in 2nd and 3rd person	1%
Paucal, Trial	not attested

Other values on pronouns, such as **paucal** or **trial**, are not encountered in SA. This is not surprising, as these values are reported to be rare outside the Pacific (Cysouw 2003: 197; Ivani 2017: 218).

Table 3 summarizes the occurence of number distinctions in independent personal pronouns in the sample languages.

Figure 1 shows the geographic distribution of the number distinctions. The following observations can be made.

First, with respect to languages with singular-dual-plural distinction in all pronouns, one can observe an areal clustering around the northern part of Peru and Ecuador/Colombia border. This areal clustering is particularly interesting since the languages in question belong to different language families (Yagua, Arabela, Murui, Miraña, Camsá and Waorani). Approximately this area (or part of the area) has been shown in the literature to have strong characteristics of a linguistic area (see Wise 2011; Valenzuela 2015, 2018). The singular-dual-plural distinction in personal pronouns has not been noted before among the shared traits of the area; however, I suggest that it can be considered as such. Besides this area, the presence of singular-dual-plural in personal pronouns is found in a few genealogically unrelated languages spoken in the Southern Cone (Mapudungun, Tehuelche and Puelche). Although the status of the Southern Cone (beyond the Chaco area) as a linguistic area is questionable, the languages spoken here do share some features (see Adelaar and Muysken 2004: 578–579; Klein 1992: 35 in Campbell and Grondona 2012: 654). However, the singular-dual-plural distinction in personal pronouns has not been noted as a shared feature. I suggest that it deserves to be included as such in further studies on areality of the Southern Cone.

Second, as for languages with singular vs. non-singular distinction in all pronouns, we find these mainly across the northern part of the continent as well as in the Chaco.

Third, languages with no number distinction in the 3rd person (but in the 1st and 2nd person) seem to cluster in Southern Amazonia and the adjacent Andean foothills. Some of these languages have been considered to be part of the Guaporé-Mamoré linguistic area (Crevels and van der Voort 2008). Whether this feature can be treated among the shared traits of the area should be examined further.

Fig. 1: Geographic distribution of number distinctions in independent personal pronouns. The map was created by George Moroz using the lingtypology package (Moroz 2017) for R.

2.2.2 Clusivity

A cross-linguistically common feature in personal pronouns is **clusivity** (Filimono-va 2005). An inclusive pronoun generally refers to the speaker + the addressee ('I and you'), whereas exclusive pronoun excludes the addressee ('I and he/she/they, but not you'). In fact, as noted by Cysouw (2003: 3) the first description of this linguistic phenomenon was done for a variety of Quechua by Domingo de Santo Tomás in his grammar of the language published in 1560.[6] In a sample of ±100 SA languages (Krasnoukhova 2014, feature NP7),[7] the parameter of clusivity in personal pronouns is found in 52 languages (52%). A slightly higher percentage, viz. 57%, is reported by Crevels and Muysken (2005), who investigate this phenomenon in 65 languages areally confined to the central-western part of SA. SA languages with the clusivity distinction show at least three different patterns (in what follows, la-bels are from Cysouw 2003). (i) 'Only-inclusive' pattern: there is a specialised form only for inclusive 'we', but no specialised form for exclusive 'we'. This is illustrated with Jaqaru (2). (ii) 'Inclusive-exclusive' pattern: there are two different forms for non-singular 1st person, i.e. 'inclusive we' and 'exclusive we', illustrated with Urari-na (3). (iii) 'Augmented' pattern: there are three different forms for non-singular 1st person, i.e. 'inclusive we', 'exclusive we', and 'unified-we' (i.e., a pronominal form which does not distinguish between different types of 'we'). Shawi in (4) ex-hibits this pattern. For an example of a language in which clusivity does not play a role in the pronominal system, see (1) above from Amarakaeri.

(2) Jaqaru (Aymaran; Coler, forthc.)

1	**na**	1 + 2	**xiwsa**
		1 + 3	**nakuna**
2	**um**	2 + 3	**umkuna**
3	**upa**	3 + 3	**upkuna**

---

6 See further Haas (1969), Hardman (1972), Manheim (1982) and Adelaar (1993), all referred to in Cysouw (2003: 3), for details on the history of the 'inclusive-exclusive' discovery.

7 Here and elsewhere, I am grateful to Simon van de Kerke, who helped coding languages for the NP domain used for the online database of grammatical properties of South American languages "South American Indigenous Language Structures (SAILS) online" (https://sails.clld.org/). The on-line database also allows to view geographic distribution of features. For example, to see the distri-bution of clusivity in independent pronouns, go to https://sails.clld.org/parameters/NP7#5/1.746/289.565

(3) Urarina (isolate; Olawsky 2006: 213)

1	**kanɨ**	1 + 2	**kana**
		1 + 3	**kanakaana**
2	**ii**	2 + 3	**ɲaara**
3	**aka**	3 + 3	**akaɨrɨ**

(4) Shawi (Cahuapanan; Hart 1988: 262 in Crevels and Muysken 2005: 318)

1	**ka**	1 + 2	**kanpo'**
		1 + 3	**kiya, këya**
		1 + 2, 1 + 3, 1 + 2 + 3	**kanpoa'**
2	**këma**	2 + 3	**kanpita**
3	**ina**	3 + 3	**inapita**

Crevels and Muysken (2005: 318–319) find the 'inclusive-exclusive' pattern in 41,5 % of the sample, the 'only-inclusive' in 12,3 %, and the 'augmented' pattern in 3,1 % of the sample languages. The inclusive category can be reconstructed as a 'basic' category (i.e. non-derived morphologically) for at least 12 of the 18 families with inclusive-exclusive distinction (Crevels and Muysken 2005: 331).[8]

As for geographic distribution of clusivity in independent pronouns, no larger areal patterns emerge (but see below for clusivity in indexation). Campbell (2012b: 292) observes that this feature "is not consistent across geographical areas […] and even within the same language family some languages can have the contrast while others lack it." However, it is of interest to note that there are some smaller areal clusters of languages with clusivity distinction. These include languages which acquired this distinction due to influences from the neighboring languages (see Crevels and Muysken 2005: 327–329). For example, the inclusive/exclusive distinction in Arawakan Mawayana was adopted from Cariban, and in Arawakan Resígaro it was borrowed from Boran (Muysken 2012: 244–255). A loss of the inclusive/exclusive distinction is also attested, as argued for Ecuadorian Quechua by Muysken (1977). Besides, clusivity in person indexation does have a geographic skewing: as shown by Birchall (2014a: 205–206), the languages of the Andes display a lower distribution of clusivity in person indexes than the rest of the continent, whereas Eastern SA as a macro-region shows a statistically significant presence of this feature.

---

8 It is of interest to note that the isolate Itonama and the Yanomamic language Sanuma have a system where the 1st person exclusive is a separate form, but the 1st person inclusive is identical to the 2nd person non-singular (Crevels field notes, in Crevels and Muysken (2005: 327), and Borgman 1990: 149) – see Daniel (2005) and Cysouw (2005) for the analyses of this kind of system from a typological perspective.

2.3 Nominal number

This section focuses on number in NPs.

2.3.1 Availability and occurrence of number

It is not surprising that SA languages with their startling genealogical diversity form a heterogeneous picture with respect to availability and occurrence of number in the nominal domain. Nevertheless, there are some tendencies.

The most common pattern with respect to presence of number in the NP involves **optional marking for all types (i.e. semantic categories) of nouns** (or NPs headed by these nouns). We find this in 39 % (36 out of 92 languages) (Krasnoukhova 2014, feature NP711). Geographically, these languages are found all across the continent; however, the Eastern Foothills, the Southwest and Southern Amazon regions stand out as hotspots of optional nominal number. For example,[9] number marking in the NP is optional in Urarina, Muniche, Northern Emberá, Arabela and Iquito, Jaqaru and Southern Aymara, Cholón, Amarakaeri, many Panoan languages (e.g. Kashibo-Katataibo, Matsés, Shipibo-Konibo, Yaminahua), Tacanan (Cavineña, Ese Ejja), Tupían (e.g. Yuqui, Mekens, Karo, Tapieté), Mamaindê, to mention just a few. While number marking is optional in general, animacy distinctions (along the animacy hierarchy), as well as topicality and specificity can play a role, as can the presence or absence of number marking on the predicate. Urarina is an example of a language in which nouns with human referents are more likely to "attract plural marking" than those lower on the animacy hierarchy, but this is far from absolute (Olawsky 2006: 369). Tapieté may provide an example of discourse factors (rather than animacy) playing a role. All types of nouns would be marked for number at the beginning of a narrative or discourse. Once these have been established as topics, number becomes optional (Hebe González, p.c., referred to in Krasnoukhova 2012: 103).

The next pattern is the direct opposite of the first: number marking is **obligatory or highly frequent on all types (i.e. semantic categories) of nouns** (or NPs headed by these nouns). We find it in 23 % of the sample (Krasnoukhova 2014, feature NP711). Geographically we find concentration of these languages in the Northwest Amazon region: in Tukanoan languages (Ecuadorian and Colombian Siona, Wanano, Cubeo, Desano), Boran (Miraña) and Witotoan (Murui), and the isolate Puinave. This type of number marking is also found in the Chaco region: in Zamucoan (Chamacoco and Ayoreo), Matacoan (Wichí, Nivaclé and Maca), and Guaicuruan (Kadiwéu). Some Andean languages show it too (Chipaya, and different Quechua

---

9 For concrete references per language, I refer the reader to Krasnoukhova (2014, feature NP711) in relation to the occurrence of nominal number discussed in this section.

varieties such as Ayacucho Quechua, Cajamarca Quechua, Inga Quechua). And it is also found in isolate Movima spoken in Bolivia.

The third pattern is when number is **obligatory (or highly frequent) only on a subset of nouns** (or NPs headed by these nouns). Animacy of the noun plays the major role here (see e.g. Lefebvre 1975 for Cuzco Quechua). So the occurrence of number follows the Animacy Hierarchy (cf. Smith-Stark 1974, Corbett 2000: 56). Another factor is topicality and/or specificity of the referent. In about 9 % of the sample, number is obligatory only on nouns with human referents. And 12 % of the sample show number on both human and non-human animates. Here we also find finer distinctions between higher animates and lower animates, which would pattern, respectively, more like humans or inanimates (see also Haspelmath 2013). In Hup, for example, plural marker is obligatory only on nouns with human referents, plural marking on non-human animates is influenced by specificity of the referent (more specific ones are usually marked, while more generic ones are not). Inanimate countable nouns are usually unmarked for number, but its presence is not ungrammatical (Epps 2008: 197). Similarly, in Teko, the occurrence of plural marking would depend on animacy, referentiality, identifiability and individualization of the referent (Rose 2011: 114). In some languages, like Trumai, a plural marker would occur on nouns with animate referents, but its use with inanimates is ungrammatical (Guirardello 1999: 56).

The last pattern, which characterizes 17 % of the SA sample, is the **absence of number marking on the noun** (or NPs headed by these nouns) (Krasnoukhova 2014, feature NP711). This is the case in Jivaroan (Aguaruna, Shuar and Wampis), Nambikwaran (Lakondê and Sabanê), Chibchan (Arhuaco), Barbacoan (Awa Pit), Arawan (Jarawara), Paéz, Chapacuran (Wari'), isolates Itonama, Kanoê, and Kwaza, among others. In these languages, nouns are unspecified for number. Languages may use different strategies when disambiguation is needed. Among the most evident ones is the use of numerals or quantifiers (e.g. 'many', 'few', etc.). Languages with no number marking on nouns at all, may nevertheless have a few stems that may differentiate number. This is the case in Jarawara (with only four nouns that have a plural form, viz. 'child', 'spirit', 'man' and 'woman', Dixon 2004: 304) and Itonama (with three nouns that have a frozen plural form, viz. 'man', 'woman' and 'girl', Crevels 2006: 163). Haspelmath (2013) noticed some "correlation with the general morphological type: Isolating languages (i.e. languages with little inflectional affixation […] appear to favor the lack or non-obligatoriness of plural marking". Among those SA languages that come closest to the isolating profile, some indeed lack number on nouns (as some Ge languages), whereas others (e.g. Trumai, Dâw) show obligatory number marking with a subset of nouns. However, most SA languages lacking number in the NP have a synthetic or even polysynthetic profile.

Irrespective of which of these four types a language belongs to, a presence of a numeral (as attributive modifier in the NP) is a separate factor. Many of the languages that have obligatory marking on all nouns (except for Quechua varieties)

Tab. 4: Occurrence of number marking on nouns (or NPs headed by these nouns).

Optional marking of number for all types (i.e. semantic categories) of nouns	39 %
Obligatory or highly frequent marking of number for all types (i.e. semantic categories) of nouns	23 %
Absent for all types of nouns	17 %
Obligatory (or highly frequent) only on a subset of nouns, specifically:	
On nouns with human referents and non-human animate referents	12 %
Only on nouns with human referents	9 %

keep their number marking when a modifying numeral is present. In languages where number is not obligatory on all types of nouns, there is some variation too. In some, number marking becomes obligatorily absent, if a numeral is present: e.g. Mosetén (Sakel 2004: 84) and Cuzco Quechua (Pieter Muysken, p.c. 2012). But more often than not, the overt expression of number simply becomes optional when a numeral is present (e.g. Imbabura Quechua, Cole 1982: 128). The issue of morphological expression of number on nouns used with a numeral modifier is discussed in Section 3.2.

Table 4 summarizes the occurrence of number marking on nouns (or NPs headed by these nouns) in the sample languages.

2.3.2 The formal marking of number

With respect to formal strategies there is a clear preference for suffixes as number markers (see Krasnoukhova 2014, feature NP710). Syntactically free number markes are much less frequent (found in Trumai, Dâw, Hixkaryana, Reyesano, Ninam, and Nhengatu). The use of plural prefixes is clearly rare, encountered only in one language (Timbira, cf. Alves 2004: 48). Reduplication as a formal device is also attested (e.g. Mosetén, Leko, Cholón, and marginally in Trumai), but it never seems to be the primary device for this function. For example, in Mosetén, plural can be marked both by a clitic and reduplication of the nominal root, but the latter expresses distributed plurality, when entities are distributed over a certain area (Sakel 2004: 83). This is in line with the arguments by Hannβ and Muysken (2014: 40) – although with a focus on Andean languages – that reduplication "often bears a particular connotation". Finally, it should be noted that number can also be conveyed using plural personal pronouns (see Goodwin Gómez [1990: 79] for animate nouns in Ninam), or so-called 'referential elements' (see Haude [2006: 150] for Movima).[10]

It is not uncommon that several markers are available for one number feature. Guaicuruan and Matacoan languages have particularly rich inventories. Not only a

---

[10] 'Referential elements' is a morphological class of its own in Movima (determined on language-internal grounds) and it comprises articles, demonstratives, and pronouns (Haude 2006: 128).

large range of values are distinguished (singular, dual, paucal, and plural – although 'plural' is described as 'collective' in some of the sources), but each of the values can be marked by several forms (see also Campbell and Grondona 2012: 645). For example, Guaicuruan Pilagá has several collective markers: one for trees and for the noun 'lake', one for humans, and yet another one for all other nouns. The latter is also used as a collective marker on mass nouns which do not take dual or paucal (Vidal 2001: 94–95).

The existence of formally different number markers distributed according to the animacy parameter is a common feature. Western and Eastern Tukanoan languages have separate forms for animate and inanimate referents (see Kotiria below). Arawakan Apurinã (in which number is generally optional) has one plural marker exclusively for humans and another with a wider distribution (including humans, animates and inanimates) (Facundes 2000: 260–262). This is also the case for Murui and Miraña. In Miraña, there are also formally different markers of number depending on a constituent it occurs on, i.e. on nouns, numerals, or relative clauses (Seifart 2005: 56).

Confirming a cross-linguistic tendency, the singular is typically left unmarked in SA languages. However, in a number of languages, the singular does receive a formal marking. Most of such languages have a rich system of classifiers that show both inflectional and derivational properties. For instance, in an Eastern Tukanoan language Kotiria, both singular and plural are marked, and the marker depends on animacy. Stenzel (2013: 113) shows that human nouns and some higher-level individual animates (see 5a) obligatorily take the suffix -ro in the singular. In the plural, human and animate nouns receive the marker -a/-~da (5b), just like inherently plural roots of collective animate nouns (referring to gregarious creatures) end in -a/ -da. Among inanimates, only count nouns are marked. Some count nouns are not marked for singular but can be pluralized by the suffix -ri; other count nouns are singularized by classifiers (as in 5c) or by -ro which functions as a partitive with inanimates (see Stenzel 2013: 130). The plural of these nouns is marked either by -ri (replacing the classifier or -ro) or, in some cases, by a special plural classifier for rounded objects (-phoka).[11] Stenzel (2013: 113) notes that in other cases (when nouns with reduced or full noun classifiers are pluralized), the classifier remains and is itself suffixed by -ri (as in 5d).

(5) Kotiria (Tukanoan; Stenzel 2013: 108, 120)
    a. *phidó-**ró***   b. *phidó-**á***
       snake-SG   snake-PL
       'snake'   'snakes'

---

11 As Stenzel (2013: 116) notes, such "irregular plural markers suggest that the language formerly had a larger paradigm of plural markers coding shape [...], and that some relic forms persist synchronically alongside the more general plural suffix -ri".

c. *hó-yó*
banana-CLF:palm
'banana tree'

d. *hó-yó-ri*
banana-CLF:palm-PL
'banana trees'

2.3.3 The semantics of number

As for semantic distinctions found in the number marking system, the most common one is the dichotomy singular vs. plural. The dual value (within NP) is found in 11 % of the sample (see Krasnoukhova 2014, feature NP720). The dual has a somewhat different distribution in these languages; specifically, its use can be limited to certain type of nouns. For example, the dual is used only with animate nouns in Ninam (Goodwin Gómez 1990: 49). In Pilagá, dual can be found with a few inanimate nouns that naturally come in pairs, for example, two feet, a pair of ribs, a pair of flip-flops (Vidal 2001: 91, 94). This confirms an observation by Acquaviva (2017), at least with respect to duals, that it is "rare to find the dual and the trial employed strictly for all and only two- and three-membered collections, without any qualifications". In Páez, dual is found with animate and inanimate nouns whose referents form certain 'constellations', but its use is not productive (Jung 2008: 132). Most examples involve human nouns producing meanings typical of dyad constructions[12] (see Evans 2006). However, we also find occurrences that are seemingly different from dyadic constructions (6c–d).

(6)   Páez (isolate, Jung 2008: 132)
      a. *p-ni:sa*
         DU-daughter
         'daughter with mother or father'

      b. *p-nej*
         DU-father
         'father with son'

      c. *p-dʲiʔp*
         DU-face
         'face to face'

      d. *p-na:sa*
         DU-person
         'shadow of a person'

The paucal is a much rarer category in SA than the dual. Languages with a marked paucal seem to be concentrated areally in the Chaco area. We find the paucal in the NP in the Guaycuruan languages Pilagá, Mocoví and Toba (but not in Kadiwéu, which has five plural suffixes, see below). The paucal as a value would generally mean 'several'. For example, in Pilagá, where dual marking is limited to only a few nouns coming in pairs (see above), a paucal marker can indicate the number "two or a few" or "between two or three and up to eight or ten" (Vidal 2001: 91). Surprisingly enough, in Pilagá there are also nouns encoding body parts normally coming

---

12 Thanks are due to Michael Daniel (p.c.) for pointing this out. Examples preserve the glossing used in the source material.

in pairs (eyes, ears, eyebrows, eyelids) that take the paucal marker (see Vidal 2001: 93–94); this may indicate that the semantic range of paucal starts indeed with 'two' in this language.[13] Outside the Chaco area, the paucal is encountered in the isolate Kwaza. The marker (-*ry*-) associated with the paucal number can be used only with nouns referring to humans and animals, and it also occurs on verbs (van der Voort 2004: 540).

The trial (referring to three items) is mainly observed in the languages in the Pacific, as noted earlier, and a few languages in Australia (see Corbett 2000: 21–22 for references). I am not aware of a SA language with a trial marker in NP.

Other categories associated with number are 'collective', 'distributive', and 'associative plural'. Corbett (2000: 111–112, 118–119) notes that the occurrences of collectives and distributives do not generally follow the Animacy Hierarchy, their use is never obligatory, and both can co-occur with a number marker. The primary function of collectives is "to specify the cohesion of a group", and thus collectives typically refer to a group of items that should be considered not individually but as a unit, often also spatially contiguous (Corbett 2000: 119, 111). Distributives, on the other hand, focus on the "separation of members of a group [...] in space, sort or time" (Corbett 2000: 111). Wichí, spoken in the Chaco region, is a language with formally different markers for collectives and distributives. As shown in (7), either of these can co-occur with a plural marker, yielding additional meanings.

(7) Wichí (Matacoan; Terraza 2009: 89, 92)
  a. *halo-**y**-**lis***
    tree-PL-DISTR
    'many trees of different sizes and forms'

  b. *hep-**ey**-**layis***
    house-PL-COL
    'many houses grouped together'

Distributive meaning is often specifically encoded by reduplication. We have already mentioned Mosetén, and this is also the case for the Andean languages Aymara, Jaqaru, Cuzco Quechua, Pukina, Kallawaya, and Uru (Hannβ and Muysken 2014: 65–66).

However, some languages do not distinguish these different number meanings, and the type of noun to which a number marker attaches can determine the semantic outcome. We see a variation even within a single family. For example, Panoan languages generally have optionally marked nominal number (Loos 1999: 236).

---

13 The numeral *dosolqa* 'two' in Pilagá is a borrowing from Spanish; however, it has been morphologically adopted and contains the morpheme -*qa* 'paucal'. This is also the number marker that occurs in agreement on nouns modified by this numeral (see Vidal 2001: 129, cited in Krasnoukhova 2012: 128).

Among them, Shipibo-Konibo has the marker *-bo* on nouns that encodes both plural and collective meaning (and it is also used to form the plural on the 3rd person pronoun, cf. Valenzuela 2003: 185). Verbs can take the plural suffix *-kan* (and *-bekon*, in case of two participants, see Section 2.4.2). In related Matsés, the marker *-bo* on human nouns can indicate "either a set of people in a group, a category of people in general, or multiple people acting separately" (8a). On nouns with non-human referents, *-bo* indicates a heterogeneous category: different kinds or types of something (8b–c). The collective reading is specified by the suffixes *-cueded* or *-beded* on verbs, either with or without the *-bo* on the NPs (Fleck 2003: 273). See Zariquiey (this volume) for yet another formal set of number markers and its distribution in Kashibo-Kakataibo, another Panoan language.

(8) Matsés (Panoan; Fleck 2003: 273, 272)
    a. *chido=bo   cho-e-c*
       woman=PL come-NPST-IND
       'A group of women are coming' / 'Women (always) come' /'Women are coming (one by one)'

    b. *poshto=bo*
       woolly.monkey=PL
       'woolly monkeys and other types of monkeys'

    c. *chompian=bo*
       shotgun=PL
       'different types of shotguns' / 'shotguns, etc.'

The category of the associative plural, argued by Corbett (2000: 105,110) to be distinct from number, generally encodes the meaning of 'X and other people associated with X', where "human proper nouns being the most common stem, and kin forming the most commonly understood associates" (Daniel and Moravcsik 2013). For South America, an associative plural marker (formally a suffix) has been suggested as one of the linguistic traits of the Vaupés-Içana Basin linguistic area (see Campbell 2012b: 307–308). Example (9) illustrates the associative marker *-ăndʼəh* in Hup, which, according to Epps (2008: 206), probably derives from *-ăn* 'OBJECT' and *-dʼəh* 'PLURAL'.

(9) Hup (Naduhup; Epps 2008: 206)
    *ʔána-**ăndʼəh** híd-ăn   gʼɔ́p-ɔ́h*
    Ana-ASS.PL   3PL-OBJ scoop-DECL
    'Ana-and-they (her children) were serving them.'

In some languages, the regular number marker can carry the 'associative plural' function. This is the case in a Tacanan language Ese Ejja: the regular plural morpheme *=kwana* has two possible meanings when attached to personal names: the referent's family or an associated group (Vuillermet 2012: 335). Mapudungun, which

lacks number marking on nouns, has a marker *-wen,* referring to a group of two or more people "who are related to each other through a relation which is indicated by the noun" (Smeets 2008: 109). For example, *foṭüm* 'son of a man' occurring with *-wen* results in 'father and son', *püñeň* 'woman's child' plus *-wen* yields 'mother and daughter', *ad* 'relative, member of a group' plus *-wen* gives 'kin' (Zúñiga 2000: 16; Smeets 2008: 109–110). These are dyadic constructions which extend beyond kin terms. Some other available examples include: *wenüy-wen* 'friends' (*wenüy* 'friend'), *peñi-wen* 'brothers' (*peñi* 'brother'), *müri-wen* 'co-wives' (from *müri* 'co-wife'), *kompaň-wen* 'partners, fellows' (from *kompaň*, a Spanish loan *compañero* 'comrade') (*idem*).

2.3.4 A typologically unusual trait: gender indexicality in number morphology

A typological feature deserving a special mention is gender indexicality in number marking. So far it is found in Kukama-Kukamiria and Omagua, two closely related languages forming a group (Omagua-Kokama) within the Tupí-Guaraní branch. These two languages have formally distinct number markers depending on the sex of the speaker, not the gender of the referent noun. Nominal plural marker in male speech is encoded by the enclitic *=kana* (in both languages), while in female speech it is marked by the enclitic *=nu* (in Kukama-Kukamiria) and *=na* (in Omagua) (Vallejos 2016: 93–94 for Kukama-Kukamiria, and O'Hagan, p.c. for Omagua). These markers reconstruct to Proto-Omagua-Kokama (O'Hagan et al. 2013, p.c.). Neither an equivalent system nor cognates are found in any other Tupían language (Zachary O'Hagan, p.c.). Much remains unknown, so gender indexicality in number morphology can be either innovation at the level of Proto-Omagua-Kokama, or a case of borrowing (of form and function) at this node, though the source language cannot (yet) be identified (*idem*). Rose's (2015) pioneering study showed that gender indexicality is much more widespread in SA than previously thought; however, gender indexicality in number morphology is rare, found only in these two languages so far (Rose 2015: 528). Examples in (10) come from Omagua. In (10a) we see the male speech plural *=kana* on the noun 'tree', whereas in (10b) the same noun 'tree' receives the female speech plural *=na*. Note that the pronominals and demonstratives in Omagua also show a gender-indexical distinction (e.g. proximals *akia* and *amai* for men and women, respectively, and distals *yuká* and *yukú*) (Zachary O'Hagan, p.c.; see also Rose 2015: 519).

(10) Omagua (Tupían; Zachary O'Hagan, p.c., speakers: Manuel Cabudivo Tuisima and Amelia Huanaquiri Tuisima)

    a. *ta=yɨ̈=pupɪ*   *ta=sakita*   *mura*   *akia*   *ɨwɨra=**kana***
         1SG:MS=axe=INSTR 1SG:MS=cut 3SG:MS DEM:PROX:MS tree=PL:MS
         'I cut it with my axe, the trees.' (Manuel Cabudivo Tuisima)

b. *iwira=**na** ukukui iwira nua=mai=**na***
tree=PL:FS fall tree be.big=REL=PL:FS
'The trees fell, the big trees.' (Amelia Huanaquiri Tuisima)

2.3.5 Nominal number as an unstable category?

Nominal number seems to be a rather unstable structural feature in SA, as genealogically related languages often show divergent behavior with respect to availability and occurrence of nominal number marking in the NP. In some families, nominal number may constitute a later development. Tupían and Chibchan languages are a case in point. In Chibchan, number "generally remains unmarked even in those languages that have the means to express it", and it is mainly restricted to human nouns and culturally relevant animals (Quesada 2007: 62). Constenla Umaña (2012: 404) suggests that "[i]t is quite possible that Proto-Chibchan did not have noun inflection", since the inflectional morphology that has been reconstructed for the proto language is all verbal. A similar argument is made for the Tupían stock. For example, Tupí-Guaraní languages lack a common plural morpheme (Jensen 1999: 151). According to Rodrigues & Cabral (2012: 524) "Proto-Tupían probably had no marker for number or gender in nouns". Individual Tupí-Guaraní languages have developed a plural marker or, in some cases, have borrowed one (Jensen 1999: 151; Gasparini 2011: 21). For instance, the plural marker *kõ* (in Wayampi) and *kom* (in Teko) are noted by Jensen (1999: 151) to be a borrowing from Cariban.

2.3.6 The presence and relevance of numeral classifiers

With respect to the relation between number, numerals and classifiers, SA is a particularly interesting terrain. In numerous languages worldwide a noun cannot occur directly with a numeral and requires the use of a classifier. Southeast Asia languages (Thai or Mandarin) are typical examples here. Nouns in classifier languages are argued to denote a concept rather than a discrete entity ('bookness' rather than 'a book', 'appleness' rather than 'an apple'), and it is the function of a classifier to create a discrete entity out of a concept, serving as an individualizer (Lyons 1977: 462 in Rijkhoff 2002: 50, *inter alia*). Greenberg (1972: 1) suggested that in a classifier language, the classifier "has the same function as a singulative does in a language with a collective-singulative distinction". One of the typological generalizations put forward in the same paper is that a language that uses a classifier in constructions with numerals would have, at most, optional plural marking (Greenberg 1972: 17, referring to Sanches 1971).

A 'prototypical' numeral classifier system would involve an inventory of classifying elements that (i) obligatorily occur in constructions with numerals and ex-

pressions of quantity, (ii) semantically categorize referents according to physical properties (shape, consistency, size and boundedness), with a possibility to assign the referent to various classes depending on the property one emphasizes, (iii) be marked only once in the NP, i.e., not being manifested in other loci in the form of agreement (neither within the NP nor on the predicate) (Dixon 1986; Grinevald 2000). An 'approximation' of this prototypical system is found in relatively few SA languages, areally confined to the Andean sphere: Tsafiki (Dickinson 2002: 57), Cholón (Adelaar and Muysken 2004: 470; Alexander-Bakkerus 2005) and extinct Mochica (Adelaar and Muysken 2004: 342). On the other hand, the lack of classifiers (as well as genders) is given as one of the typological traits of Andean languages (see Dixon and Aikhenvald 1999: 8, 10; Adelaar 2008: 31). The three languages are curious in this respect. Furthermore, there are a few languages outside the Andean sphere with the use of classifiers specifically in the context of quantification, namely, Chibchan languages Cuna and Chimila (Adelaar and Muysken 2004: 66, 79) and a few North Arawakan languages (Aikhenvald 1999a: 83). While I have just introduced a prototypical numeral classifier system and mentioned their representatives in SA, this prototypical system is rather an exception than the rule in SA. Specifically, this deviation from a prototypical classifier system has been argued for languages spoken in the Northwest Amazon (see Payne 1987; Derbyshire and Payne 1990; Aikhenvald 1999b; Seifart and Payne 2007) as well as Southwest Amazon regions (see van der Voort 2004: 179, 2005). What appears to be more widespread in SA is a classifier system of a 'multifunctional' type (Krasnoukhova 2012: 204–217). Although there is some variation in the degree to which they are manifested (even between closely related languages), the properties of a 'multifunctional' classifier system typically combine four functions: (i) semantic categorization (adding semantic content to the head noun), (ii) morphological derivation (deriving new stems from either verbal or noun roots or stems), and/or creating full NPs, when occurring on a modifying constituent; (iii) reference tracking in the discourse, and (iv) to a lesser degree, syntactic function of agreement. While some languages use formally different sets of classifiers in different constructions, in the majority of languages we see a largely identical set of classifiers occurring in different syntactic environments: e.g. on head nouns, modifying constituents, question words, on verbal and non-verbal predicates (see also Aikhenvald 2000: 204).

Crucially, when a classifier occurs on a numeral (or other modifiers such as demonstratives and property words) it often has both derivational and inflectional properties (Seifart and Payne 2007: 383). Thus a classifier does not only serve semantic categorization (i.e. creating an entity out of a concept expressed by the head noun), it functions as a derivational marker on roots used as modifiers. Since the same classifer typically occurs also on other constituents within the same NP (e.g. either on other modifiers and/or head noun), it functions as an agreement marker. Due to these properties, we encounter the pattern [numeral + classifier + noun + plural], which, as I have just indicated, has been argued to be cross-linguistically

rare. According to Sanches and Slobin 1973 (referred to in Rijkhoff 2002: 29) there are few or possibly no languages "in which the noun must take a plural marker while the attributive numeral combines with a true classifier". However, we see this pattern in SA, for example in Baure (11).

(11) Baure (Arawakan; Danielsen 2007: 171)
      po-no   *mapi-**no***   *eton-**anev***
      other-CLF:human two-CLF:human woman-H.PL
      'the other two women'

Similarly, (12) from Miraña shows the classifier *-ʔí* 'bunch' on the numeral *ma:kíní-* 'three', while the noun, with the same classifier in agreement, also takes a plural marker.

(12) Miraña (Boran; Seifart 2005: 131)
      *ma:kíní-**ʔi**-ßa*   (*úhɨ-**ʔi-:nɛ***)
      three-CLF:bunch-PL banana-CLF:bunch-PL
      'three bunches of banana'

Example (5) above from Kotiria is yet another illustration. Thus a typologically unusual pattern of co-occurrence of number marking and classifiers turns out not to be so unusual in SA.

2.4 Verbal number

2.4.1 Preliminaries

Verbal number is a term commonly used to refer to the category of number in relation to events.[14] Adopting Corbett's (2000: 246) terminology, two types of verbal number are distinguished here: (i) 'event number', referring to non-singular occurrence of an event or action encoded by the verbal predicate, and (ii) 'participant number', referring to non-singular number of participants involved in a state or action encoded by the verbal predicate. Crucially, to qualify for either type of verbal number, the formal marking must be on the verb. Example (13) from Teko shows that the event of 'going out' happened more than once. Specifically, disyllabic redu-

---

14 'Pluractionality' is another common term for the phenomenon. I refer the reader to Cabredo (2010) and Mattiola (2019) who address the terminology and main arguments in the key studies on verbal number cross-linguistically. See also Corbett (2000: 243).

plication of the verb indicates a repetitive action, in this particular example, of one and the same participant (Rose 2005: 353).[15]

(13) Teko (Tupían; Rose 2005: 353)
õhẽ-õ-hem-ne   *o-ʔa*
RED-3-go.out-CONTRAST 3-fall
'He goes out again and falls down.'

Example (14) from Yurakaré illustrates a stem alternation of the verbal predicate, signaling participant number. The language has less than a dozen of such verb pairs (van Gijn 2006: 191). In these pairs, stems of intransitive verbs differ for the number of the subject, while stems of transitive verbs for the number of the object. In other words, intransitive verbs indicate plurality of the subject, while transitive verbs – plurality of the object.

(14) Yurakaré (isolate; van Gijn 2006: 192)
   a. **dele**-Ø   *ti-biskotcho*
      fall.SG-3 1SG-bread
      'My bread fell.'

   b. **ñeta**-Ø   *ti-biskotcho*
      fall.PL-3 1SG-bread
      'My breads fell.'

The phenomenon of verbal number has been in focus for a few individual SA languages (see Henry 1948: 198 and D'Angelis 2004 for Kaingáng, Rose 2005, 2007 for Teko and Tupí-Guaraní in general, Crevels 2006 for Itonama, Müller & Sanchez-Mendes 2008 for Karitiâna, Salanova 2014 for Mẽbengokre, and Mattiola & Gildea (subm) for Akawaio and Cariban in general). And it is also addressed – albeit to a different degree – in grammatical descriptions of some languages that have it. Mueller (2013) is a typological study of 'Tense, Aspect, Modality and Evidentiality' in SA languages and is relevant here as it includes (among others) an analysis of iterativity and habituality, i.e. aspectual categories closely associated with event number. Furthermore, there is a recent typological study by Mattiola (2019) on pluractionality, based on a world-wide sample, including 29 SA languages. As far as I am aware, there is no comparative study on both types of verbal number specifically for SA yet. In this respect, the present section aims to give a first exploratory account of these categories focusing on this part of the world. The analysis is based on data from 70 SA languages representing 41 genealogies: 31 language families with more

---

15 Reduplication in Teko shows two patterns, each with a different semantics of verbal number (see Rose 2005).

than two members and 10 isolates (see Table 5). Before we turn to the discussion of the data, I briefly specify the scope of the present exploration.

Event number includes such aspect categories as 'iterative' and 'distributive'. 'Iterative' refers to the marking of actions "that are repeated instantly" regardless of the number of times and generally has following characteristics: the same subject carries out the action, the action is not customary, there is no change in meaning of the verbal action,[16] and the repetitions happen immediately after each other ('event-internal' repetition [Cusic 1981: 61]) (Mueller 2013: 97). 'Distributive' refers to actions taking place with certain intervals in time ('event-external' repetition [Cusic 1981: 61]) and space, and with no strict regularity involved. There is yet another aspectual category – 'habitual', which refers to an action that is customary, i.e. "repeated/or be in a state over a certain amount of time" (Mueller 2013: 93; Comrie 1976: 27–28). While habituals can be associated with event number, they do not belong to the core functions of verbal number (Mattiola 2019: 57, ch: 2). I leave habituals outside the scope of this study.[17]

As for participant number, this type of verbal number can be hard to distinguish from *nominal* number manifested on the verb (agreement). Corbett (2000: 252) and Mattiola (2019: 86) discuss the problem of differentiation of nominal and verbal number, acknowledging a possibility of indeterminacy. Nevertheless, Corbett (2000: 252–255 building on Durie 1986) suggests some diagnostics for verbal number based on a variety of languages for which sufficient data allow teasing nominal and verbal number apart. These diagnostics[18] include: (i) a formal marking of (participant) number on the verb operates on the S(ubject)/P(atient)-oriented basis,[19] viz. the marking indicates (and is controlled by) the number of the S-argument for intransitives and the P-argument for transitive verbs. (ii) The values available for verbal participant number are often different from those available for (pro-)nominal number. Semantically, in general, the primary function of verbal number of this type is noted "not to enumerate entities, but to quantify the effect of actions, states, and events" (Mithun 1988: 214 in Corbett 2000: 248).

Nevertheless, there are many cases in the SA data where a differentiation is not evident. SA languages frequently index arguments on the predicate (Birchall 2014a:

---

16 According to Mueller (2013: 99), some verbs may change their meaning when receiving a marker for repetition. She exemplifies it with Kukama-Kukamiria: the verb *itika* 'throw' has a different meaning when suffixed with the iterative suffix *-ka* (which on other verbs marks repetition, see Mueller 2013: 103), namely *itika-ka* 'to separate, get divorced' (Vallejos 2010: 359 in Mueller 2013: 99).

17 For habituals in SA I refer the reader to Mueller (2013). Note that to ensure uniformity in judgement for the whole sample all data used for the present analysis have been collected by myself from grammatical descriptions.

18 While these are suggested by Corbett (2000: 252) as "diagnostics", one may rather regard these as generalizations emerging from data.

19 Corbett (2000: 252) used the term 'ergative basis' in his discussion. I refer to it as 'S/P-oriented basis'. The latter term is due to Michael Daniel (p.c.).

45, 222, 232). Indexation can involve person markers morphologically conflated with number (and in some cases gender), but it can also be done by markers that encode only number (with no person or gender information) (see Section 3.1). In some of these languages, such number markers are obligatory on the verbal predicate, but in many others they are optional. An extra dimension to the issue is given by the fact that some languages share its number markers between nouns and verbs, i.e. formally the same number morphology can be used (often optionally) on either verbal predicate or the NP.[20]

Verbal markers that encode only number (but not person or gender) are subsumed in Table 5 under the 'participant plurality' type of verbal number. However, the occurrences which do *not* follow the S/P-oriented basis (and thus can be only S/A-oriented or only P-oriented) are highlighted with shaded cells. The latter could be regarded as argument indexing by number morphology only (see Section 3.1). A question remains whether it makes sense – for some SA data at least – to tease apart nominal number and verbal number of the participant plurality type.

2.4.2 Overview of the results: Verbal number in SA

Data on verbal number in the sample languages are presented in Table 5. The first vertical section shows whether the language has a so-called 'mixed' type of verbal number, i.e. where one and the same formal device can encode both event plurality and participant plurality (Corbett 2000: 249). The second vertical section indicates whether a language has a formal device that encodes event plurality only. Finally, the third vertical section shows if a language has a formal device that encodes exclusively participant plurality. As noted earlier, number markers which do not operate on the S/P-oriented basis are indicated by shaded cells. It must be noted that it is not always evident from the available sources whether a certain strategy is 'limited' to only one type of verbal number, or whether it can encode both. Thus the section with 'mixed' verbal number only includes cases for which positive evidence is found in the source. It is possible that certain instances of either event or participant number will be reclassified as the 'mixed' category when more data are available.

As for formal marking of verbal number, three main strategies are found in SA (illustrated below): (i) reduplication of the verb stem, (ii) stem alternation, and (iii) affixes/enclitics. This echoes what Mithun (1988) describes for North American languages, and corresponds to the most common means cross-linguistically (see Matti-

---

20 For example, plural markers shared by nouns and verbs have been noted for Boruca and Cuna (Chibchan) by Quesada (2007: 63). This also the case in Teko (Tupian) and Trinitario (Arawakan) (Françoise Rose, p.c. 2020), isolate Leko (van de Kerke 2009: 306) and Tsafiki (Barbacoan) (Dickinson 2002: 57).

ola 2019: ch.3). It is worthwhile to note that in one language in our data (Itonama), verbal classifiers are among the means to encode participant plurality (see Table 5).

As can be seen in Table 5, encoding of either event or participant number (in the senses outlined above) is found in the overwhelming majority of SA genealogies surveyed (viz. at least in 35 out of 41 surveyed). In 14 genealogies (out of 41) we find a 'mixed' type of verbal number, i.e. where one formal strategy is confimed to encode both event and participant number. These genealogies are: Araucanian, Aymaran, Cariban, Enlhet-Enenlhet, Harakmbut, Macro-Ge, Matacoan, Panoan, Peba-Yagua, Tupían, Yanomaman, and the isolates Itonama, Movima and Warao. In the two large stocks, Macro-Ge and Tupían, both types of verbal number are found in closely and more distantly related languages, suggesting that the presence of this linguistic feature goes back to an earlier stage of these two language families. For example, singular and plural stems of positional, motion and *dicendi/faciendi* verbs may go to Proto-Tupían (Rodrigues and Cabral 2012: 527). Furthermore, reduplication of stems as a means to convey either event and/or participant number is deep-rooted in Tupían languages (Rodrigues 1953; Rodrigues and Cabral 2012: 539).

In the sample, only few genera showed little or no evidence for verbal number. For example, in Ayoreo (Zamucoan), verbal morphology in general is very limited, and event quantification can be rendered only by lexical means, such as 'again' adverbial (Bertinetto 2009, p.c.). In Iquito (Zaparoan), there are a few derivational verbal morphemes that encode event quantification, but they are rare in frequency and synchronically unproductive (Lai 2009: 515–521). In some languages (Barbacoan, Chocoan, and isolates Camsá and Leko), it is uncertain if we deal with verbal number or rather with nominal number (indicated with shaded cells in Table 5).

The geographic distribution of the sample languages with 'mixed' verbal number, event number or participant number is shown in Figure 2. Geographically, we find the languages with the 'mixed' type of verbal number in different regions of SA. Languages with exclusively event number are concentrated in Western and Southern Amazon, the adjacent eastern foothills and the Central Andes. Languages with exclusively participant number are concentrated along the western part of SA corresponding to the Andean sphere, but are also found in Southern Amazon (particularly the northeastern part of Bolivia-Brazil border), parts of central Brazil, and in the Chaco area.

The geographic distribution of each formal strategy is shown in Figure 3–5 and discussed later in the section.

Tab. 5: Mixed verbal number, event and participant number in SA.
(*Conventions:* + *indicates that a phenomenon is present, ± indicates that it is marginal and unproductive. Morphological number markers are provided, unless their inventory is too large, in which case a '+' is used instead. Cells are left empty if no evidence for a phenomenon is found. Shaded cells indicate markers that do not follow the S/P-oriented basis*).

Family	Language + iso-code	'Mixed' verbal number			Event number			Participant number			Reference
		Red.	Stem	Morph.	Red.	Stem	Morph.	Red.	Stem	Morph.	
Araucanian	Mapudungun [arn]	+		-tu							Zúñiga (2000: 46, 55); Smeets (2008: 271–272)
Arawakan	Apurinã [apu]						-poko				Facundes (2000: 309)
	Baure [brg]						-porely			-he	Danielsen (2007: 284, 234)
	Parecís [pab]						-hitiya				Brandao (2014: 304)
	Tariana [tae]						-nipe				Aikhenvald (2003: 329)
Arawan	Jarawara [jaa]			-ra	+		-rima		+		Dixon (2004: 278, 162, 74, 543–545)
	Kulina [cul]							+		-bakhi	Dienst (2014: 105, 133–134)
	Paumarí [pad]				+		-bakhia		+		Chapman&Derbyshire (1991: 280–281)
Aymaran	Central Aymara [ayr]			-ra							Hardman (2001: 77)
	Southern Aymara [ayc]			-ra						-rpaya	Coler (2014: 368–371, 373)
	Jaqaru [jqr]				+					-rqaya	Hardman (2000: 52, 86–87)
Barbacoan	Awa Pit [kwi]									-a, -na	Curnow (1997: 66, 75)
	Tsafiki [cof]									-la	Dickinson (2002: 228)

Tab. 5 (continued)

Family	Language + iso-code	'Mixed' verbal number			Event number			Participant number			Reference
		Red.	Stem	Morph.	Red.	Stem	Morph.	Red.	Stem	Morph.	
Boran	Miraña [boa]						+				Thiesen & Weber (2012: 138)
Cariban	Hixkaryana [hix]										Derbyshire (1985)
	Panare [pbh]						-pëti				Payne & Payne (2013: 185)
	Tiriyó [tri]	+									Meira (1999: 36); Mattiola and Gildea (subm)
Chicham	Aguaruna [agr]				+		-kawa,-ha		+	-ina,-aha	Overall (2017: 122, 331, 254, 234)
	Wampis [hub]				+		-kaua			-ina,-ara	Peña (2015: 834, 211)
Chapacuran	Wari' [pav]				+			+	+		Everett & Kern (1997: 337–339)
Chibchan	Ika [arh]							+ (?)			Frank (1990: 56)
	Border Cuna [kvn]							+ (?)	± (?)	-mal(a)	Forster (2011: 162, 167)
Chocoan	Embera [emp]									-da	Mortensen (1999: 9, 41)
Chonan	Tehuelche [teh]						-e				Fernandez Garay 1998: 151)
Enlhet-Enenlhet	Sanapaná [spn]			et-							Van Gysel (2017: 44, ft.14)
Guaicuruan	Kadiwéu [kbc]									-waji	Sandalo (1997: 51)
	Mocoví [moc]									-lo/-er	Grondona (1998: 147)
	Pilagá [plg]						- tapiñi			-a/-tol-lo	Vidal (2001: 262, 162)

Family	Language		-m(b)a~mã				Van linden (forthc.)
Amarakaeri	Amarakaeri [amr]		-m(b)a~mã				Van linden (forthc.)
Macro-Ge	Rikbaktsa [rkb]	+			+		Silva (2011: 170–172, 105, ft.43)
	Apinajé [apn]				+		Oliveira (2005: 60, 128–129)
	Mẽbengokre [txu]	+	+				Salanova (2014: 267)
	Akwẽ-Xerente [xer]				+		Filho (2007: 89)
	Kaingang [kgp]	+	+				Cavalcante (1987: 57–65, 81), D'Angelis (2004: 76, 80)
	Xokleng [xok]	+			+		Urban (1985: 176–177, ft.4), Gakran (2005: 38–40)
Matacoan	Nivaclé [cag]			ʃaʔne~ -xaʔne, -tʃʔe~ -kʔe		-k(l), -s	Fabre (2016: 293, p.c)
	Wichí [mtp]		-li/ -wito			-hen	Terraza (2009: 157–160)
Naduhup	Hup [jup]		±				Epps (2008: 579–582)
Nambiquaran	Mamaindê [wmd]			-tɨk			Eberhard (2009: 406)
Panoan	Kashibo-Kakataibo [cbr]			-bëkin, -rat~rakët		-kan	Zariquiey (2011: 398, 424, 439)
	Matsés [mcf]	+	-ne, -ban		+		Fleck (2003: 338–339, 360, 380, 453)
	Shipibo-Konibo [shp]				±	-bekon, -kan	Valenzuela (2003: 81, 521, 594–596)

Tab. 5 (continued)

Family	Language + iso-code	'Mixed' verbal number			Event number			Participant number			Reference
		Red.	Stem	Morph.	Red.	Stem	Morph.	Red.	Stem	Morph.	
Peba-Yagua	Yagua [yad]			-tííty			-jqq, -jayqq				Payne and Payne (1990: 393–396)
Quechuan	Huallaga Quechua [qub]				+		-(y)kacha, -cha			+	Weber (1989: 150, 321, 143–144)
Takanan	Araona [aro]				+						Emkow (2006: 586–587)
	Cavineña [cav]				+		-(ne)ni				Guillaume 2008: 349, 206)
Tukanoan	Cubeo [cub]				+						Chacon (2012: 207–208)
	Ecuadorian Siona [sey]				±						Bruil (2014: 226)
Tupían	Karitiâna [ktn]	+	+								Storto (2014: 409, 423)
	Karo [arr]				+				+		Gabas Jr. (1999: 58, 238)
	Mekens [skf]				+				+	-kwa	Galucio (2001: 55–56, 104)
	Awetí [awe]	+									Drude (2014: 185)
	Nheengatu [yrl]	+									Da Cruz (2014: 128)
	Kukama-Kukamiria [cod]				+		-ka(ka)				Vallejos (2016: 208, 212, 219–220)
Uru-Chipaya	Chipaya [cap]				+						Cerrón-Palomino (2006: 164)
Witotoan	Murui [huu]				+					-zi (?)	Wojtylak (2017: 336, 260, ft.283)
Yanomaman	Yanomami [wca]			=(r)a	±						Perri Ferreira (2017: 326–327, 258)

								Reference
Zamucoan	Ayoreo [ayo]						+	Bertinetto (2009: 24–25, p.c.)
Zaparoan	Iquito [iqu]		±					Lai (2009: 515–521)
isolate	Camsá [kbh]	-ang, -at						O'Brien (2018: 89–90)
isolate	Itonama [ito]	+	+		-he, -te/-tye	-ke, -'cha'ke	+	Crevels (2006, 2012: 268–269)
isolate	Kwaza [xwa]		+ (?)	±			+	van der Voort (2004: 384–387, 457, 562–563)
isolate	Leko [lec]	-aya						van de Kerke (2009: 306)
isolate	Mosetén-Chimané [cas]				-dyi, '-		+	Sakel (2004: 267, 270)
isolate	Movima [mzp]					-ka		Haude (2006: 433–434)
isolate	Páez [pbb]						+	Jung (2008: 65)
isolate	Trumai [tpy]						+	Guirardello (2014: 220)
isolate	Urarina [ura]	+					+	Olawsky (2006: 418–424, 408–410)
isolate	Warao [wba]					-bu		Romero-Figeroa (1997: 99)
isolate	Yurakaré [yuz]		+		i-/y-, -uma		+	van Gijn (2006: 189–190, 191–192)
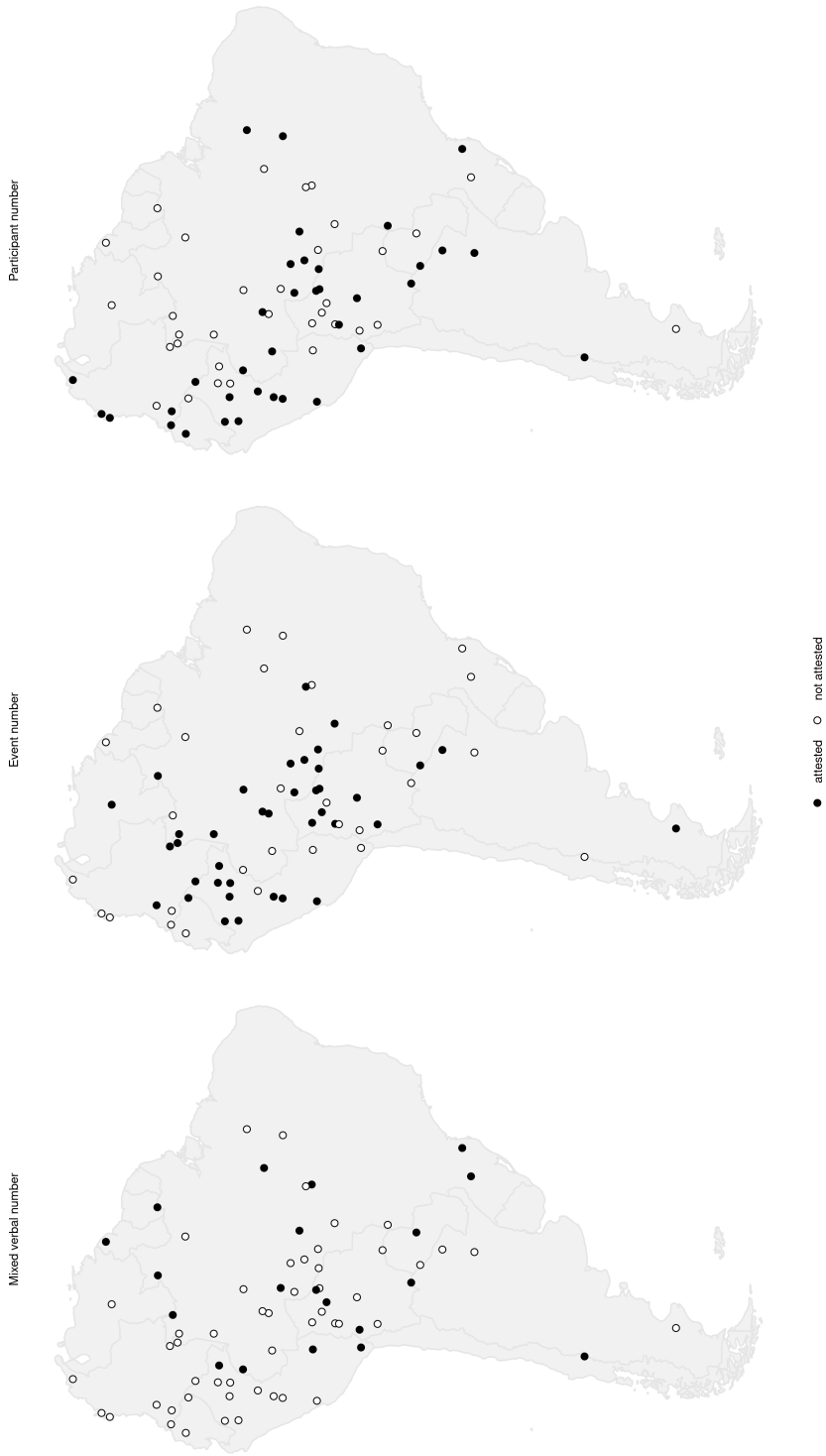
Fig. 2: Geographic distribution of verbal number in the sample: languages with a 'mixed' type of verbal number, languages with only event number, and languages with only participant number. The map was created by George Moroz using the lingtypology package (Moroz 2017) for R.

As noted above, there are three main formal means to encode verbal number: (i) reduplication of the verb stem, (ii) stem alternation, and (iii) affixes/enclitics. Each of the means is illustrated next.

The Tupían languages Juruna and Nheengatu demonstrate reduplication that serves to mark event number (15) (see also 13 above from Teko), and both event and participant number in (16). In Teko reduplication shows two patterns, each with different semantics of verbal number (see Rose 2005).

(15) Juruna (Tupían; Fargetti 2001: 181, cited in Rodrigues and Cabral 2012: 540)
    a. *Ídja   iyâ   awi*
       woman water drink
       'the woman drank water'

    b. *Ídja   iyâ   **awi-wi***
       woman water drink-RED
       'the woman drank water more than one time'

(16) Nheengatu (Tupían; Da Cruz 2014: 130)
    *ta-**yuká-yuká***
    3PL.A-kill-RED
    'They killed [a large number of animals]!'

Stem alternations are illustrated by Karitiâna (17), Akwẽ-Xerente (18), and Yurakaré (14). In Karitiâna stem alternations are available not only for lexical verbs, but also for the copula 'be' and for (imperfective) auxiliaries (Storto 2014: 411–413). The plural form of the imperfective auxiliary in (17b) signals more than one 'lifting' event, and the number of participants can be one or more. This concerns the S-participant for intransitive verbs and P-participant for transitive verbs. Plural forms of some lexical verbs (e.g. ones with an irreversible effect) entail an interpretation of more than one participant involved (Storto 2014: 409, p.c.). We saw this also in Nheengatu (16). It should be noted that Karitiâna has no number marking on nouns otherwise (see Storto 2014: 402).

(17) Karitiâna (Tupían; Storto 2014: 413)
    a. *Ø-py-mangat **tyka**-dn   taso Luciana*
       3-ASSERT-lift IPFV-NFUT man Luciana
       'Luciana is lifting a/the man/men.'

    b. *Ø-py-mangat **agi**-dn   taso Luciana*
       3-ASSERT-lift IPFV.PL-NFUT man Luciana
       'Luciana is lifting a/the man/men (more than once).'

In Akwẽ-Xerente, stem alternations signal number of participants involved, which can be one, two or more than two.

(18) Akwẽ-Xerente (Macro-Ge, Central Ge; Filho 2007: 89)

    a. *kuba mɜ̃t*   *ø-**wara***
       boat 3.PST.RFV.REAL 3-leave.SG
       'The boat left.'

    b. *kuba mɜ̃t*   *ti-**ssamrõ***
       boat 3.PST.RFV.REAL 3-leave.DU
       'The boats (they both) left.'

    c. *kuba mɜ̃t*   *ø-**ssakre***
       boat 3.PST.RFV.REAL 3-leave.PL
       'The boats (all) left.'

Event and participant number encoded with morphological markers can be illustrated with Yanomami (19) and Amarakaeri (20). In Yanomami, the distributive/non-punctual marker *=ra/=a* indicates that an event took place several times or that multiple entities are involved. The meaning of a verb on which the marker occurs determines a specific reading of the verbal number (Perri Ferreira 2017: 326–327).

(19) Yanomami (Yanomaman; Perri Ferreira 2017: 327)

    a. *kaxa*   *pë=riyë* *wa=**a**=ki=ma*
       caterpillar 3PL=raw eat=DISTR=PFV2=PST
       'He ate the caterpillars.'

    b. *ĩhĩ=tëhë kõa*   *ãyo=kɨ=tixiti=**ra**=ri=ma*
       ANA=PRS firewood CLF:forewood=PL=cut_in_pieces=DISTR-PFV1=PST
       'Then [she] chopped the firewood in pieces [one after the other].'

In Amarakaeri, nouns are unmarked for number. Plurality of participants is indicated on the verb by means of the prefix *mba-~ma-~mã-* (Van linden, forthc.). The marker can also express event plurality.

(20) Amarakaeri (Harakmbut; Van linden forthc.)

    Herman o-n-a   *Bernardo-ta*   *me-niŋ-**mba**-kaʔ'-e-Ø*
    Herman 3SG.ind-spat-say Bernardo-ACC 2/3SG>1SG-ben-vpl-make-ITER-2IMP
    pĩã
    arrow
    'Herman has arrows made by Bernardo.' (Lit. 'Herman says to Bernardo: "Make arrows for me!"')

In the isolate Itonama, there are several means to signal event and participant number (see Crevels 2006). For instance, there is a suffix *-'ke* which "usually denotes event plurality" but "may denote participant plurality, especially when event plurality is already indicated by another strategy" (Crevels 2006: 167). The most typologi-

cally interesting strategy to encode participant number is a set of classifiers on verbs. Among 17 verbal classifiers that categorize their referent according to animacy, posture, and physical properties (shape, consistency, size), there are five pairs of suppletive forms that additionally encode number (singular and plural) (Crevels 2012: 268). Example (21) shows one existential verb carrying the classifier -*pu* for 'lying, singular' referents and another verb taking the classifier -*di* for 'sitting, plural' referents.

(21) Itonama (unclassified; Crevels 2012: 268)

*o-si-**pu***	*u-k'a'ne*	*opi*	*wa'ihna*	*o-si-**di***	*a-chïpa*
DV-EX-CLF:lying.SG	DV-one	fish	DM	DV-EX-CLF:sitting.PL	DV-two

u-puwe
DV-bird
'There is one fish and two birds.'

A formal strategy may have either event number or participant number as its primary, or preferred, reading and another reading as its semantic extension. As adumbrated already, pluralization of events encoded by verbs with certain semantics (e.g. those with an irreversible effect) can push the interpretation of plurality of a participant. Consider Karitiâna. Recall that all NPs in Karitiâna are unmarked for number. Thus both nouns ('man' and 'potato') in (22a) can be interpreted as either singular or plural. However, a reduplicated verb root can never have a reading of a singular event (Storto 2014: 406). The verb form in (22b) thus encodes that an event of 'peeling' took place more than once, and the semantics of the verb forces the interpretation of a plural object (Storto 2014: 406).

(22) Karitiâna (Tupían; Storto 2014: 408)

a. *taso* *Ø-na-**ohok**-a-t* *ohy*
man 3-DECL-peel-STMF-NFUT potato
'A/the man/men peeled a/the/some potato(es).'

b. *taso* *Ø-na-**ohok~ohok**-‹o›na-t* *ohy*
man 3-DECL-peel.PL-STMF-NFUT potato
'A/the man/men peeled potatoes (more than once).'

Similarly, a plural form of a verb such as to 'catch' (formally marked by stem alternation) entails non-singular number of an affected participant (23b). Example (23c) is ungrammatical as it involves a plural form of the verb for 'catch' and an overtly singular affected participant. A clear difference due to semantics of the verb can be seen as compared to the example (17b) above, where the plural form of the predicate signals more than one lifting events, but the number of participants can be one or more.

(23) Karitiâna (Tupían; Storto 2014: 410)

    a. *myhin-t / sypom-p   taso Ø-naka-**ot**-Ø   myhin-t 'ip*
       one-POSP / two-POSP man 3-DECL-catch-NFUT one-POSP fish
       'One/two man/men caught one fish.'

    b. *myhin-t / sypom-p   taso Ø-naka-**piit**-Ø   'ip*
       one-POSP / two-POSP man 3-DECL-catch.PL-NFUT fish
       'One/two man/men caught fish (more than once).'

    c. **myhin-t / sypom-p   taso Ø-naka-**piit**-Ø   myhin-t   'ip*
       one-POSP / two-POSP man 3-DECL-catch.PL-NFUT one-POSP fish
       'One/two man/men caught one fish (more than once).'

In a discussion of verbal number in Balinese, Arka and Dalrymple (2017: 300–305) point out that a reading – either that of a plural event or a plural participant – can depend on a number of factors, such as the type of verb (e.g. event or state), its concrete semantics, the animacy of a participant, and whether it is referred to with a proper noun or a common noun. This observation is valid for quite a few SA languages too (e.g. see Rose 2005: 357 for Tupí-Guaraní and Dietrich 2014: 306 for Tupí in general, Perri Ferreira 2017: 326 for Yanomami, among others).

    There is another comparative feature worth mentioning about the SA data. Arka and Dalrymple (2017: 294–295) discuss modification of pluractionality by numerals in Balinese, noting the argument by Xrakovskij (1997) that cross-linguistically such structures are unacceptable and dispreferred. Just like in Balinese, we see numeral modification of pluractional predicates in Karitiâna (24). In this language, interpretation of event plurality is possible not only for a significant number of events, but also for two events, as in (24) (see Storto 2014: 407).

(24) Karitiâna (Tupían; Storto 2014: 407)

    sypom-p *Ø-na-**pon~pon**-Ø   João sojxaty   kyn*
    two-POSP 3-DECL-shoot.PL-NFUT João wild.boar POSP
    'João shot twice at a/the boar(s).'

As can be seen in Table 5, both event number and participant number can be marked in a language by more than one strategy. For example, participant number in Kulina and Rikbaktsa can be indicated by reduplication and by stem alternations. Event number in Mẽbengokre and Kaingang is encoded by reduplication and by prefixing morphology. Inventories of verbs that show stem alternations differ in size: with only few verbs in some languages (e.g. Tupían, see Rodrigues and Cabral 2012: 510), around 20 in some other (e.g. Jarawara, see Dixon 2004: 543, and Kaingang, see Cavalcante 1987: 65), and many more in others (e.g. Wari', see Everett and Kern 1997: 377). Semantics of these verbs typically include position, motion and directional verbs (e.g. 'come', 'go'). It seems that, typically, stem alternation, redu-

plication, and affixal morphology – if present – will not be available to the same set of verbs but would rather occur in 'complementary distribution'.[21] However, this type of information is seldom present in grammatical descriptions, so no clear conclusions can be drawn as of now. Although verbal and nominal number involve somewhat different semantics, they can complement each other: for example, in Akwẽ-Xerente, in contexts involving verbs that do not show stem alternations and thus cannot encode participant plurality, nominals receive the suffix -*nõrĩ*[22] marking non-singular (Filho 2007: 90).

As for formal means to encode event number, reduplication and morphological affixes are equally common in the sample, and some languages have both means. As seen from Table 5 (second section), reduplication is used in at least 21 languages (and in three additional languages it is unproductive), and morphological markers are employed in at least 22 languages (and in one additional language it is unproductive). It is worth noting that no language in the sample uses stem alternations to encode exclusively event number,[23] with no participant plurality reading involved (as far as one can conclude from the available data).

As for participant number, less than five sample languages use reduplication to encode exclusively participant number (with no event plurality readings involved). In contrast, stem alternations to express exclusively participant number are found in at least 13 sample languages. Morphological markers encoding only participant number are found in 26 sample languages. 17 of these (identified with shaded cells) do not follow the S/P-oriented basis, suggested in the literature as a diagnostic feature of the participant plurality type of verbal number. These might be regarded as instances of nominal number instead, although number in the nominal domain can be rather scarce in these languages. For example, in Tsafiki, number is generally optional. In the nominal domain, the plural marker is more likely to occur with human nouns than with non-humans. It can occur on both the noun and the predicate, or only on the noun, or only on the predicate (Dickinson 2002: 57). The formal marker is the same (-*la*). (25) illustrates plural marking on the verb. It is of interest to note that the marker -*la* is also used to derive plural forms of personal pronouns (Dickinson 2002: 65). Note that while the noun *unila* 'man' ends in '*la*' (which might historically involve the plural marker), synchronically, the noun does not seem to be plural, evidenced from the fact that examples available with the noun taking the plural marker (*unila-la* 'men', see Dickinson 2002: 246). We return to the issue of verbal vs. nominal number on the verb in Section 3.1 below.

---

21 For example, see Storto (2014: 409) for Karitiana and Cavalcante (1987: 65) for Kaingang.

22 This suffix is also used to form plural forms for all personal pronouns (see Filho 2007: 117).

23 This is in line with results in Mueller (2014, feature TAME3-6) who reports no languages that use 'suppletion' as a strategy for iterative marking.

(25) Tsafiki (Barbacoan; Dickinson 2002: 57)

tsan-ke-to=bi,   *unila mantiminni jelen=chi*   *ji-**la**-i-e*
SMBL-do:VCL-SR=LOC man EMPH   jungle=LOC go-PL-become:VCL-DECL
'When they had done this the men went to the jungle.'

It is of interest to consider the geographic distribution of each formal strategy.

Figure 3 illustrates the distribution of languages that use reduplication for encoding 'mixed' verbal number, event number and participant number. As can be seen, reduplication for encoding 'mixed' verbal number is scattered across the continent, but seems to 'avoid' the western part of the continent. Reduplication for encoding exclusively event number favors the Southern Amazon region and parts of Western Amazon plus the adjacent eastern foothills of the Andes. Reduplication for exclusively participant number is rare, as noted above, found in three unrelated languages in Southern Amazon (Wari, Kwaza and Kulina) and two Chibchan languages spoken in the North of SA. Figure 4 also includes a map called 'not attested', which shows the languages where reduplication as means for marking verbal number is not attested. It can be seen that reduplication as a strategy for verbal number is primarily absent in the Chaco area. Furthermore, languages spoken in the North of Ecuador and parts of Northern Amazonia seem to lack it too.

Figure 4 shows the distribution of affixation as a strategy for encoding verbal number. Similarly to reduplication, affixation for encoding 'mixed' verbal number is scattered across SA. Affixation for encoding only event number is found in the Central Andes and along the foothills, as well as in parts of Western and Southern Amazon. The distribution of languages that use affixation to encode exclusively participant number is more striking: the languages are concentrated along the western edge of the continent corresponding to the Northern and Central Andes, as well as the eastern foothills. Another area that stands out for using this strategy is the Chaco area. Finally, affixation for any type of verbal number seems to be absent in the central part of Brazil, as well as parts of Northern Amazonia.

Figure 5, finally, shows the distribution of stem alternation as means to encode verbal number. Recall that no language in my sample was found to use stem alternation to encode exclusively event number. Stem alternations for a 'mixed' verbal number are found only in two languages (Kaingang and Karitiâna) – these are shown in Figure 5 with a name tag. All the other languages shown in the Figure use stem alternations to encode exclusively participant number. Geographically they are confined to the Southwest Amazon region and the northern part of Peru. The strategy is also found in several languages of the Macro-Je family, spoken in Brazil.

Another interesting question relates to number values available in the verbal domain for participant number. A two-fold distinction, i.e. singular vs. non-singular (or plural), is the most common one in the sample. However, several languages additionally distinguish a dual: Kulina, Jarawara, Akwẽ-Xerente, and Shipibo-Konibo. One language, Urarina, additionally has a paucal category. In Shipibo-Konibo, the dual is marked by the suffix *-bekon*: it signals that the action is carried out by two
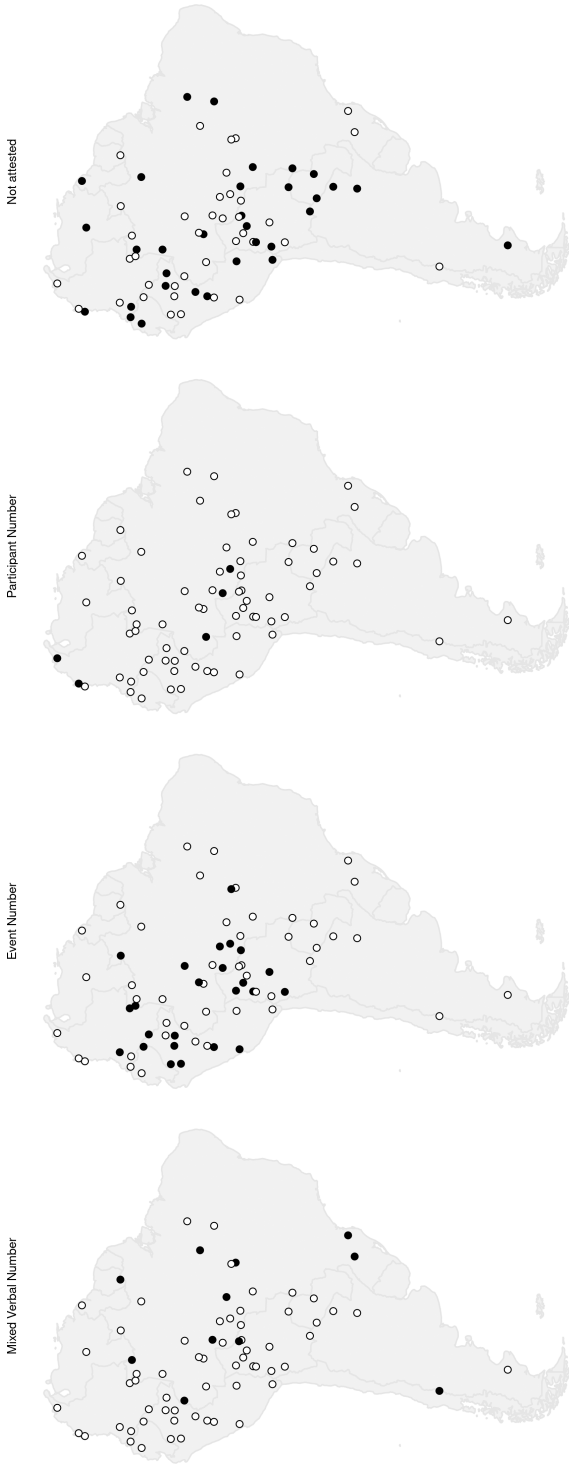
Fig. 3: Geographic distribution of *reduplication* as a strategy to encode verbal number. The map was created by George Moroz using the lingtypology package (Moroz 2017) for R.
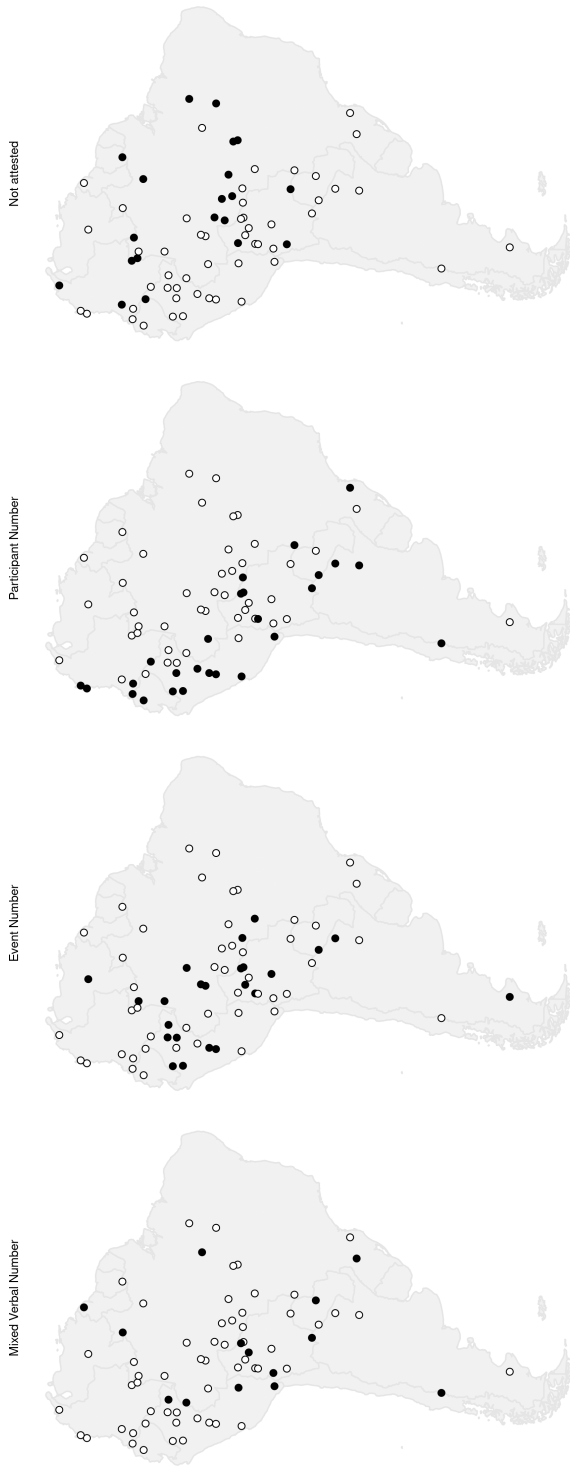
Fig. 4: Geographic distribution of *affixation* as a strategy to encode verbal number. The map was created by George Moroz using the lingtypology package (Moroz 2017) for R.
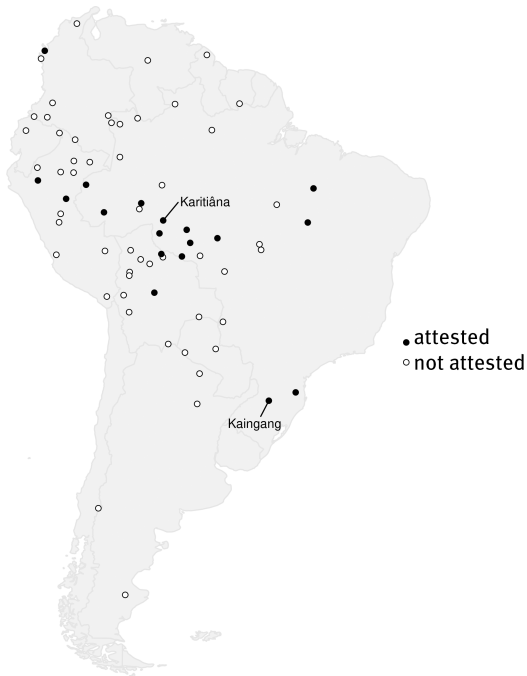
Fig. 5: Geographic distribution of *stem alternation* as a strategy to encode verbal number. The map was created by George Moroz using the lingtypology package (Moroz 2017) for R.

participants (Valenzuela 2003: 281). Kulina, Jarawara, and Akwẽ-Xerente show verb stem alternations for singular-dual-plural. We saw these distinctions in Akwẽ-Xerente in (18) above. Example (26) from Kulina illustrates the dual form of the verb 'move'. The intransitive verb form indicates the dual number of the subject, in this case semantically agreeing with the NP containing a numeral (formally a verb 'be two'). The verb 'move' has the following stem forms: *kha-* for singular, *weda-* for dual (3rd person), *kada-* dual (1st + 2nd person), and *hawi ha-* (where *ha-* is an auxiliary verb) for plural (Dienst 2014: 134).

(26) Kulina (Arawan; Dienst 2014: 139)
      moto   Ø-ka-pame-e   Ø-ka-**weda**-hona-ni
      boat(F) 3-CLF-be.two-REL.F 3-CLF-move.DU-HITHER-DECL.F
      'Two boats are coming.'

Kulina, additionally, can mark dual by reduplication. (27a) features an intransitive verb, a stative verb in this case, indicating dual number of the subject, and (27b) shows a transitive verb, encoding dual number of the object.

(27) Kulina (Arawan; Dienst 2014: 144, 106)
  a. **weshe~she** *ti-na-ni*
     clean.DU   2-AUX-DECL.F
     'You two are clean.'

  b. *takara* **dama~ma** *o-na-na*
     chicken grab.DU   1SG-aux-ifut
     'I'm going to grab two chickens.'

In related Jarawara, there are about ten stance verbs with stem alternations, distinguishing singular, dual, and plural of the subject participant. Some have interesting neutralizations. The verb for 'be hanging from a hook, lie in a hammock' has the same form (*-wina-*) for a singular and plural subject, but a different form (*-wata-*) for a dual subject (see Dixon 2004: 544, also for a possible motivation). Then the verbs for 'sit' and 'stand' have different forms for singular (*-ita-* and *-wa(a)-*, respectively); however, their semantics is neutralized for dual and plural subject. Thus, the forms *joro -na-* or *teme -na-* (in free variation) can mean both 'sit' or 'stand' (for dual subject) and the forms *-naho* (for an animate) and *sii-na-* (for an inanimate) is both 'sit' or 'stand' (for plural subject).

   Urarina is interesting as it has sets of productive derivational[24] suffixes for position, color and shape verbs. These suffixes can encode up to five number distinctions: singular, dual, paucal, greater plural, and distributive plural. These are argued to derive different types of verbs for which different number values will be available. For example, derived intransitive states (e.g. 'it is lying on its side') can distinguish singular, dual, paucal, greater plural, and distributive plural. Derived intransitive action verbs (e.g. 'it is lying down on its side') distinguish between singular, paucal, and greater plural (thus, no dual). Derived transitive actions (e.g. 'he laid something on its side') distinguish only singular and plural (Olawsky 2006: 379–380, 408). As Olawsky (2006: 408–409) shows, the paucal in Urarina refers to 'a few', which may comprise from three and up to five participants depending on the context. Besides, there is also some speaker variation for the interpretation of 'few'. Greater plural refers to 'many' (four or more), but should be part of one 'group' gathered at the same place, which differs from distributive plural which refers to 'various groups' or collectives present at different places (*idem*).

2.5 Number value asymmetries across domains

We find a number of asymmetries with respect to number values across domains.

   As for the nominal domain: among the languages that distinguish between singular-dual-plural in pronouns (see Table 2), Kashibo-Kakataibo, Waorani, Malayo

---

24 See Olawsky (2006: 379–384) for more details.

and Chimila do not have the dual distinction for nouns. Mapudungun and Puelche do not mark number on nouns at all. More number distinctions in nouns than in pronouns are found in Pilagá. While pronouns make only singular vs. non-singular distinction, some nouns can also occur with the dual and the paucal marker (Vidal 2001: 91).

Particularly interesting is that a few languages systematically encode more number distinctions in verbs than in nouns or pronouns. This is the case for Kulina, Jarawara, Shipibo-Konibo, Akwẽ-Xerente, and Urarina. As discussed in Section 2.4.2, these languages encode participant number on verbs, differentiating singular, dual, and plural number, and there is an additional paucal distinction in Urarina. All except Jarawara have a singular vs. non-singular distinction available for pronouns and for nouns. In Jarawara, this distinction is found in pronouns, but nouns do not differentiate number (except for four nouns 'child', 'spirit', 'woman' and 'man', cf. Dixon 2004: 289, 304–305).

The presence of dual as well as paucal values marked in the verb is curious from a cross-linguistic perspective. Furthermore, these SA languages challenge Corbett's (2000: 251) generalization that nominal number systems typically show more number values than verbal number systems.

3 Agreement and the syntax of number

'Agreement' can be defined as a "systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978: 610 in Corbett's 2006: 4). In the SA data, one can find cases of 'agreement' ranging from more to least canonical ones.[25] In what follows I reserve the term 'agreement' for the nominal domain, and use 'indexing' for any indication of an argument's properties on the verb.

3.1 Indexing on the verb

SA languages frequently index arguments on the verbal predicate (Birchall 2014a: 45, 222, 232). Indexing is done by bound person markers. Birchall (2014a: 45) finds

---

25 As Corbett (2006) argues, agreement can show properties which range from more canonical to less canonical ones. Specifically, obligatory marking is more canonical, whereas optional marking is less canonical. Inflectional morphology (affixes) is more canonical than suppletive forms. Productive marking of agreement (e.g. every verb shows agreement) is more canonical than sporadic marking (e.g. only some verbs agree). Agreement is more canonical if it appears on a target irrespective of the presence of the controller, and less canonical if it appears only if the controller is absent. A target agreeing with only one potential controller is more canonical then if a target has a choice of controllers (Corbett 2006: 14–18).

that "[t]he presence of a full person-based verbal argument marking system of at least one argument role in basic main clause constructions occurs in 66 of the 74 languages in the sample, with 42 of these languages displaying only indexation of arguments and no core case marking". Bound person markers can conflate information on person with number and/or gender. Birchall (2014b, features ARGEX2-2, 2–5, 2–8) examines the way argument marking is realized for three types of argument: S, A, and P.[26] Number is fused with person in about half of the sample languages (viz. 53 % in case of S-argument, 51 % in case of P-argument, and 49 % in case of A-argument).[27] As already noted in Section 2.4.2, some SA languages can index a referent with a marker that encodes only number but not person or gender. In Birchall (2014b: *idem*), indexation by markers encoding only number is found in 31 % for S, 30 % for A, and 18 % for P.

Teasing apart nominal number and verbal number of participant plurality type is not always straightforward. Table 5 (third column, shaded cells) included cases which can be seen as argument indexing by a 'dedicated' number marker. For example, in Northern Embera (28), the subject noun can be unmarked for number, unless "large numbers are being talked about" (Mortensen 1999: 41). The verbal predicate, however, obligatorily carries plural marking when S is plural (*idem*).

(28) Northern Embera (Chocoan; Mortensen 1999: 62, cited in Birchall 2014a: 51)
      *hãŭ čapa-ra   wã-pi-shi-**da**-a*
      that brother-ABS go-CAUS-PST-PL-DECL
      'They made that guy leave.'

In Southern Aymara, in a sentence with a plural subject, neither the noun needs to carry the plural marker (-*naka*), nor does the verb need to take the plural marker (-*pha*) (Coler 2014: 166, 425). Optional occurrence of number marker on either the noun or the verb was also shown for Tsafiki (25), where nominals and verbs use the same plural maker (-*la*). In related Awa Pit, nouns are always unmarked (Curnow 1997: 131), but verbs can take an optional number marker when the referent is human. For a plural Subject, the number marker is -*a*, while for a plural Object, the marker is -*na* (Curnow 1997: 66, 75). When both Subject and Object are human and plural, the choice for a plural marker depends on the participant the speaker wishes to focus on (Curnow 1997: 183). This is shown by a comparison of (29a) where the focus is on the Subject, and (29b) where it is on the Object (*idem*).

---

26 The abbreviations stand for: S = single participant of an intransitive clause, A = agent-like participant of a transitive clause, P = patient-like participant of a transitive clause.
27 The percentage is calculated based on data in Birchall (2014b, features ARGEX2-2, 2–5, 2–8).

(29) Awa Pit (Barbacoan; Curnow 1997: 183)

    a. *uspa=na   au=mɨza   kal   ki-wayn-**a**-mtu-y*
       3PL(NOM)=TOP 1PL=(1/2PL)ACC work do-AUXIL-PL:SUBJ-IPFV-NONLOCUT
       'They are helping us work.'

    b. *uspa=na   au=mɨza   tɨt   ta-**na**-tɨ-zi*
       3PL(NOM)=TOP 1PL=(1/2PL)ACC cord give-PL:OBJ-PST-NONLOCUT
       'They passed us a cord.'

Finally, there are also languages in which number information is not encoded on the verb at all. This is the case for 31 % for P, 21 % for A, and 16 % for S (see Birchall 2014b).

3.2 Agreement in the NP

With a demonstrative as modifier, agreement in number occurs in less than 30 % of the languages. With adjectives (defined semantically[28]), it is found in only 23 % (see Krasnoukhova 2014, features NP231 and NP211). Languages vary with respect to obligatoriness, conditions on the occurrence (e.g. animacy of the head), and the locus of the number marker (i.e. only on modifier, or both on head and modifier). Plural forms of demonstratives can be derived by a nominal number marker or have a separate plural form. Although I do not have consistent data on adjectives, some languages, like Tupían Gavião and Arawakan Tariana, have suppletive plural forms for some adjectives. In Tariana, this includes the adjective *hanu* 'big, wide, long', which can be used with singular referents only, and the form *male* 'thick, heavy', which is used as the plural term, but can also occur with a singular referent (Aikhenvald 2003: 173). In Gavião, there are several adjectives with a suppletive plural form: *póòk* 'big.SG' and *tą́k* 'big.PL', *cíciít* 'small.SG' and *cî̀k* 'small.PL', and others. These plural stems in Gavião can only modify nominals that are plural (Moore 1984: 145–146).

    As for NPs with a numeral functioning as attributive modifier, the following observations can be made. It should be mentioned that the presence of a numeral can influence the occurrence of a number marker on the noun – this issie was briefly discussed in Section 2.3.1. Here I focus on agreement in number which can be marked on a numeral functioning as attributive modifier. A brief remark on numerals is in order.[29] The majority of SA languages have native numerals for lower num-

---

[28] In some languages property words form a separate part of speech, whereas in others property words may align with nouns or with verbs.

[29] Morphosyntactic properties of numerals of SA languages are discussed in Krasnoukhova (2012: 111–117).

bers only. These can be compound expressions, etymologically transparent (e.g. Chapacuran Wari': *xica' pe* 'to be alone' for 'one' and *tucu caracan* 'to face each other' for 'two', cf. Everett and Kern 1997: 338, see also Aikhenvald 2012: 350–359). Elaborate systems, with a possibility to count up to high numbers with a certain systematicity are relatively rare. These are mainly, but not exclusively,[30] found in the Andean languages (e.g. Quechua, Aymara) (cf. Adelaar 2008: 24).

Agreement in number which is manifested on an attributive numeral (and the head noun in some languages) is found in 11,5 % (10 out of 86 languages) (Krasnoukhova 2014, features NP221). Here, again, it can be influenced by the properties of the noun (with animacy as the relevant factor), or by those of the numeral (e.g. obligatory agreement with numerals 'one' and 'two' and optional for higher numerals).

Example (30) from Puinave illustrates agreement on the modifying numeral and the adjective. In Puinave, plural marking occurs on all nouns, but number agreement on a numeral seems to be confined to human and animate nouns only (available examples show no agreement with inanimates). Note that both semantic numerals and adjectives are encoded by verbs in Puinave which need to be nominalized by prefix *i-* for attributive use (Girón 2008: 372).

(30) Puinave (unclassified; Girón 2008: 386)
 *i-pɤ́i-**ot**   yót-**ot,** i-bɤk-noḰ-**ot***
 ATTR-three-PL dog-PL ATTR-be.white-ASP-PL
 'these three white dogs …'

Miraña is a language with obligatory number agreement only for the dual distinction. Example (31) shows agreement with the dual suffix with the numeral for 'two'. Starting from numeral 'three', morphological agreement is optional, and if present, it is encoded by the 'plural' marker.

(31) Miraña (Boran; Seifart 2005: 130)
 *mí-ʔí-**:ku**   (úhɨ-ʔí-**:ku**)*
 two-SCM:bunch-DU (banana-SCM:bunch-DU)
 'two banana bunches'

In Pilagá, the paucal morpheme *-qa* occurs as agreement marker with the numeral 'two'. As discussed in Section 2.3.3, only few nouns can be marked for dual in Pilagá, and the semantic range of the paucal maker can start at 'two' (cf. Vidal 2001: 91)

---

30 Boran Miraña spoken in the Northwest Amazon region is a notable exception: it has a quinary system with native numerals running until 400 (Seifart 2005: 130).

(32) Pilagá (Guaycuruan; Vidal 2001: 129)

qanač'e na'   tayiñi **dosol-qa** *na'   **emek-qa***
CONJ   CLF:PROX south two-PAUC CLF:PROX house-PAUC
'And in direction to the south, there are two houses.'

In some SA languages the attributive use of numerals within NP is not possible. These are used either as verbal predicates or as adverbs modifying predicates. The former situation is found in Cariban languages, Karo and Kamaiurá (Tupí), and Sabanê (Nambikwaran). The latter – in Jarawara (Dixon 2004: 559) and Wari' (Everett and Kern 1997: 338). In Wari', the numeral for 'two' *tucu caracan* (lit. 'to face each other') is replaced by Portuguese *dois* nowadays, but as Everett and Kern show, the borrowed numerals (like *oito* for 'eight' in (33b)) adopt the syntactic behavior of intransitive verbs.

(33) Wari' (Chapacuran; Everett and Kern 1997: 339)

a. ***tocu-u   caracan   na   xirim***
   face-1SG each.other 3SG:rp/p house
   'There are two houses.' (Lit.: 'The houses face each other')

b. ***Oito   na   cawaxi'   nucun***
   Eight 3SG:rp/p dry.season POS:3SG.m
   'He is eight years old.' (Lit.: His years are eight.)

With respect to word order, a numeral preceding the head noun is the most common word order in SA languages, if the numeral occurs attributively at all. In languages where numerals can both precede and follow the noun, the most common factor determining the order is the pragmatic status of the referent, and less expectedly, animacy. For example, (in)definiteness of the referent influences the word order in the Chibchan languages (Frank 1990: 31), some Tukanoan languages, and Arawakan Tariana. As for animacy, in Mosetén, there is a tendency for numerals to occur before an inanimate noun, but after an animate one (Sakel 2004: 82). The properties of the numeral can also play a role. For example, in Itonama, the native numerals 'one' and 'two' can only occur before the noun (see 16 above), while borrowed Spanish numerals can precede or follow the noun (Crevels 2012, p.c.). The situation is the reverse in Panoan Shipibo-Konibo: native numerals 'one' and 'two' show free word order, while borrowed numerals beyond 'two' are of Quechua origin and obligatorily precede the noun as in Quechua (Valenzuela 2003: 235).

4 Semantics and discourse

As I do not have systematic data on pragmatic functions of number marking, I limit the discussion to mentioning a few relevant issues.

Some SA indigenous societies show avoidance behavior associated with respect towards 'in-laws'. Linguistically this is reflected in a special 'avoidance' speech register, which influences lexicon and grammar. This cultural trait is best known from Australian Aboriginal societies (e.g. Haviland 1979), but also from parts of Africa and Southeast Asia. Aikhenvald (2012: 360–364) gives a nice sketch of some SA languages with an avoidance speech style. Relevant to us is that the obligatory use of plural forms of pronouns when addressing or referring to a singular individual is one of the linguistic features of the avoidance style. This is the case in Kamaiurá (Tupían). Seki (2000: 326–327) points that among individuals involved in avoidance relations, there is a prohibition of pronouncing the name of the relative and people involved need to be addressed or referred to using 2nd plural forms (34a) or the 3rd plural form (which is a demonstrative with the plural suffix *-wan*) (34b). Sentences in (34) can thus be used in reference to a single person (in 'avoidance' speech) or to several people in usual speech.

(34) Kamaiurá (Tupían; Seki 2000: 327)
    a. *ko   katy rak **pe**-ko*
       Farm DIR TAM 2PL-COP
       'You were at the farm'

    b. ***a'e=wan**-a*   *o-kwahap*
       DEM.DIST-PL-NUCLEAR.CASE 3-know
       'They know.'

Plural forms can be associated with polite speech and respect, and may expand beyond the avoidance register. In the isolate Urarina, the verbal suffix *-tɕe* occurs as part of argument markers on verbs in the context of 2nd person, marking plural. As Olawsky (2006: 356) notes it "also functions as a politeness marker in conversations between opposite sex and specific in-law relationships". This suffix is homophonous with a politeness marker "which occurs irrespective of person, and can attach to introducers, 3rd person verb forms and neutral verb forms" (p.357).

However, it is not always the case that plural pronouns are associated with respect, as languages may have other means to this purpose.

Furthermore, we have an interesting case of plural marking in the discourse tracking system in Yagua (Peba-Yagua). As Payne and Payne (1990: 377) discuss, fully-specified NPs are used only in marked contexts and are generally avoided. They note that one strategy to avoid fully-specified NPs "is to treat one plural participant as singular when two groups are interacting. In such cases, the most topical or in some sense 'salient' group is treated properly in terms of its semantic plurality. The other group is treated as singular." Payne and Payne give an example that "if adults and children are interacting, the adults will be referred to as plural and the children as singular. If humans are interacting with animals, the humans will be plural and the animals singular. If 'good guys' are interacting with 'bad guys', it is

predictably the 'good guys' who are treated as plural while the 'bad guys' are treated as singular" (p. 377). There is yet another strategy visible in the participant reference system: the use of 2nd and 3rd person dual forms in reference to women who have had children (which also extends to the kinship system, with special terms for kins who have had children) (see Payne and Payne 1990: 377–378).

5 Conclusion

There is substantial variation among SA languages as to the availability and overt realization of the category of number.

As noted in Section 2.2, 1st and 2nd person independent pronouns (i.e. speech act participants) distinguish number categories in the majority of SA languages. 3rd person pronouns are less often marked for number. Quite a few SA languages derive non-singular forms of all personal pronouns by adding a plural maker to the singular pronouns, a feature noted to be typologically rare. About 17 % of SA languages encode dual in independent pronouns. I have suggested that these show areal clustering around the Northern part of Peru and the Ecuador/Colombia border, as well as in the Southern Cone (Figure 1). Finally, there are also languages in SA in which the category of number in pronouns is not an obligatory distinction.

In Section 2.3, it was shown that nominal number in the NP is obligatory only in 23 % of the languages. Occurrence of number marking can be optional (conditioned by several factors, animacy being the most common one), which is the case for 21 % of the languages. Interesting is that number is always optional in about 39 % of the languages, and in another 17 % number in NP is simply unavailable. Among typologically unusual traits observed in the domain of number we saw gender indexicality. As discussed in Section 2.3.4, the Tupían languages Kukama-Kukamiria and Omagua have number suffixes which differ depending on the gender of the speaker. Yet another rare pattern cross-linguistically concerns a co-occurrence of a plural marker with a classifier (Section 2.3.6), which is typical of SA languages with rich classifier systems of the 'multifunctional type'.

Section 2.4 presented results of a first exploratory survey of verbal number in SA. Following Corbett (2000: 246), I distinguished event number and participant number as subtypes of verbal number. Event and/or participant number is a widespread phenomenon in SA: we find it in at least 35 of the 41 genera's surveyed. Geographically, sample languages with the 'mixed' type of verbal number are found in different regions of SA. Languages with exclusively event number are concentrated in Western and Southern Amazon, the adjacent eastern foothills and the Central Andes. Languages with exclusively participant number are concentrated along the western part of SA corresponding to the Andean sphere, but are also found in Southern Amazon (particularly in the northeastern part of Bolivia-Brazil border), parts of central Brazil, and in the Chaco area (see Figure 2). I also considered the

geographic distribution of each formal strategy (Figure 3–5). Conclusions included the following points. Stem alternations as a formal strategy to encode verbal number seems to be geographically confined to the Southwest Amazon region and the northern part of Peru. The strategy is also found in several languages of the Macro-Je family, spoken in Brazil. As for reduplication as strategy, it can be noted that reduplication for exclusively event number shows a geographic patterning with a concentration in the Southern Amazon region, parts of Western Amazon plus the adjacent eastern foothills of the Andes. Reduplication for any type of verbal number turns out to be largely absent in the Chaco area. As for affixation as formal means, the geographic distribution of the languages that use affixation for exclusively participant number shows an areal pattern: these languages are concentrated in the Chaco area, as well as along the western edge of the continent corresponding to the Northern and Central Andes and the eastern foothills.

Finally, I pointed out that teasing apart verbal number of the participant plurality type and nominal number manifested on the verb is an interesting challenge for SA, with many in-between cases. Further research, including on the diachrony of number in nominal and verbal domains, may bring new insights.

In Section 3, the question of agreement in number was addressed. Agreement in the NP is present but is not particularly common in the sample: it is found in less than 30 % with demonstratives as modifiers, in 23 % with adjectives (defined semantically) and in about 10 % with numerals (defined semantically). As for the verbal domain, it is common for SA to index arguments on the verb by means of person markers (Birchall 2014a, 2014b). About half of the sample languages conflate person and number information, and about one third of the sample languages can index a referent by number markers only (*idem*).

Section 4 briefly addressed the use of plural markers in discourse tracking, as well as in speech related to avoidance register and politeness.

Acknowledgements

Abbreviations

1	1st person
2	2nd person
3	3rd person
A	agent
ABS	absolutive
ACC	accusative
ANA	anaphoric pronoun
ASP	aspect marker
ASS	associative
ASSERT	assertative mood
ATTR	attributivizer
AUX	auxiliary
AUXIL	auxiliative
BEN	beneficiary/benefactive
CAUS	causative
CLF	classifier
COL	collective
CONJ	conjunction
CONTRAST	contrastive
COP	copula
DECL	declarative
DEM	demonstrative
DIR	directional
DIST	distal
DISTR	distributive
DU	dual
DV	dummy vowel
EMPH	emphatic
EX	existential
F	feminine
FS	female speech
H	human
IFUT	immediate future
IMP	imperative

IND	indicative
INSTR	instrumental
IPFV	imperfective
ITER	iterative
LOC	locative
M	masculine
MS	male speech
NFUT	non-future
NOM	nominative
NONLOCUT	non-locutor person marker
NPST	non-past
OBJ	object
PAUC	paucal
PFV	perfective
PL	plural
POS	possessive
POSP	postposition
PROX	proximal
PRS	present
PST	past
REAL	realis
RED	reduplication
REL	relative
RP/P	realis past/present
SG	singular
SMBL	semblative
SPAT	spatial
SR	same referent
STMF	stem formative
SUBJ	subject
TAM	tense/aspect/mood
TOP	topic marker
VCL	verbal class marker
VPL	verbal plural

References

Acquaviva, Paolo. 2017. Number in language. Oxford Research Encyclopedias. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.61 https://oxfordre.com/

Adelaar, Willem. 1993. La categoría del plural inclusivo en las lenguas americanasinterpretación y realidad. In Maxim P. A. M. Kerkhof, Hugo de Schepper & Otto Zwartjes (eds.), *España: ¿Ruptuta 1492?*, 207–216. (*Diálogos hispánicos* 11). Amsterdam: Rodopi.

Adelaar, Willem. 2008. Towards a typological profile of the Andean languages. In Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof (eds.), *Evidence and counter-evidence. Essays in honour of Frederik Kortlandt. Volume 2: General linguistics*, 23–33. Amsterdam: Rodopi.

Adelaar, Willem & Pieter Muysken. 2004. *The languages of the Andes*. Cambridge: Cambridge University Press.

Aguilera, Oscar. 2001. *Gramática de la lengua Kawésqar.* Temuco: Corporación de Desarrollo Indígena.

Aikhenvald, Alexandra. 1999a. The Arawak language family. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The Amazonian Languages,* 65–105. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra. 1999b. Areal diffusion and language contact in the Içana-Vaupés basin, North-West Amazonia. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The Amazonian Languages*, 385–416. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra. 2000. *Classifiers. A typology of noun categorization devices.* Oxford: Oxford University Press.

Aikhenvald, Alexandra. 2003. *A Grammar of Tariana*. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra. 2012. *The Languages of the Amazon*. Oxford: Oxford University Press.

Alexander-Bakkerus, Astrid. 2005. *Eighteenth-Century Cholón*. Utrecht: LOT.

Alves, Flávia de Castro. 2004. *O timbira falado pelos Canela Apaniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê*. Unicamp: Universidade Estadual de Campinas dissertation.

Araujo, Gabriel Antunes de. 2004. *A grammar of Sabanê: a Nambikwaran language*. Utrecht: LOT.

Arka, I Wayan & Mary Dalrymple. 2017. Nominal, pronominal, and verbal number in Balinese. *Linguistic Typology* 21(2). 261–331.

Bacelar, Laércio Nora 2004. *Gramática da língua Kanoê*. Katholieke Universiteit Nijmegen dissertation

Bertinetto, Pier Marco. 2009. Ayoreo (Zamuco): A grammatical sketch. (Quaderni del Laboratorio di Linguistica della, N.S., 8.) Pisa: Scuola Normale Superiore di Pisa.

Birchall, Joshua. 2014a. *Argument marking patterns in South American languages*. Utrecht: LOT Dissertation Series 352.

Birchall, Joshua. 2014b. Argument marking (ARGEX). In Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad. *South American Indigenous Language Structures (SAILS) Online*. Leipzig: Online Publication of the Max Planck Institute for Evolutionary Anthropology. (Available at http://sails.clld.org).

Brandão, Ana Paula Barros. 2014. *A reference grammar of Paresi-Haliti (Arawak)*. University of Texas at Austin dissertation.

Bruil, Martine. 2014. *Clause-typing and evidentiality in Ecuadorian Siona*. Utrecht: LOT Dissertation Series.

Cabredo Hofherr, Patricia. 2010. Verbal plurality and event plurality. Course materials for the Summer School on Linguistic Typology, 15 August – 3 September 2010, Leipzig. https://www.eva.mpg.de/lingua/conference/2010_summerschool/pdf/course_materials/Cabredo2010plV.pdf

Campbell, Lyle. 2012a. Classification of the indigenous languages of South America. In Campbell, Lyle & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 59–166. Berlin: Mouton de Gruyter.

Campbell, Lyle. 2012b. Typological characteristics of South American indigenous languages. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 259–330. Berlin: Mouton de Gruyter.

Campbell, Lyle. 2017. Language isolates and their history. In Lyle Campbell (ed.), *Language isolates*, 1–18. London & New York: Routledge.

Campbell, Lyle & Verónica Grondona. 2012. Languages of the Chaco and Southern Cone. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 625–667. Berlin: Mouton de Gruyter.

Cavalcante, Marita Pôrto. 1987. *Fonologia e morfologia da língua kaingáng: o dialeto de São Paulo comparado com o do Paran*. University of Campinas dissertation.

Cerrón-Palomino, Rodolfo. 2006. *El Chipaya o Lengua de los Hombres del Agua*. Lima: Fondo Editorial, Pontificia Universidad Católica del Perú.

Chacon, Thiago Costa. 2012. *The phonology and morphology of Kubeo: The documentation, theory, and description of an Amazonian language*. University of Hawai'i at Mānoa dissertation.

Chapman, Shirley & Desmond C. Derbyshire. 1991. Paumarí. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, 161–354. Berlin: Mouton de Gruyter.

Clairis, Christos. 1985. *El qawesqar: lingüística fueguina, teoría y descripción*. (Estudios Filológicos, Anejo, 12.) Valdivia: Universidad Austral de Chile.

Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*, 151–179. Amsterdam: John Benjamins.

Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North-Holland.

Coler, Matt. 2014. *A Grammar of Muylaq' Aymara: Aymara as spoken in Southern Peru*. Leiden: Brill.

Coler, Matt. Forthcoming. Jaqaru. Matthias Urban (ed.) *The Oxford Guide to the Languages of the Central Andes.* Oxford: Oxford University Press.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Constenla Umaña, Adolfo. 2012. Chibchan languages. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 391–440. Berlin: Mouton de Gruyter.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville. 2006. *Agreement*. Cambridge: Cambridge University Press.

Crevels, Mily. 2006. Verbal Number in Itonama. In Grażyna Rowicka & Eithne Carlin (eds.), *What's in a Verb?,* 159–170. LOT: Utrecht University.

Crevels, Mily. 2012. Itonama. In Mily Crevels & Pieter Muysken (eds.), *Ambito Andino,* 233–294. La Paz: Plural Editores.

Crevels, Mily & Pieter Muysken. 2005. Inclusive-exclusive distinction in the languages of central-western South America. In Elena Filimonova (ed.), *Clusivity. Typology and case studies of the inclusive-exclusive distinction*, 313–339. (Typological Studies in Language 63). Amsterdam: John Benjamins Publishing Company.

Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*, 151–179. Amsterdam: John Benjamins.

Curnow, Timothy. 1997. *A grammar of Awa Pit (Cuaiquer): An indigenous language of south-western Colombia*. Australian National University dissertation.

Cusic, David. 1981. *Verbal Plurality and Aspect*. Stanford University dissertation.

Cysouw, Michael 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

Cysouw, Michael. 2013. Inclusive/Exclusive Distinction in Independent Pronouns. In Matthew Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/39, Accessed on 2018 10 09.)

D'Angelis, Wilmar da Rocha. 2004. Concordância verbal de número em Kaingang: Algumas pistas. *LIAMES* 4. 71–81.

Da Cruz, Aline. 2014. Reduplication in Nheengatu. In Gale Goodwin Gómez & Hein van der Voort (eds.), *Reduplication in Indigenous Languages of South America*, 115–141. Leiden: Brill.

Daniel, Michael & Edith Moravcsik. 2013. The Associative Plural. In Matthew Dryer & Martin Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck

Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/36, Accessed on 2020 04 06.)

Daniel, Michael. 2005. Understanding inclusives. In Elena Filimonova (ed.), *Clusivity. Typology and case studies of the inclusive-exclusive distinction*. [Typological Studies in Language, 63], 3–48. Amsterdam: John Benjamins Publishing Company.

Daniel, Michael. 2013. Plurality in Independent Personal Pronouns. In Matthew Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/35, Accessed on 2018 09 30.)

Danielsen, Swintha. 2007. *Baure. An Arawak language of Bolivia* (Indigenous Languages of Latin America 6). Leiden: CNWS Publications.

Derbyshire, Desmond & Doris Payne. 1990. Noun classification systems of Amazonian languages. In Doris Payne (ed.), *Amazonian linguistics. Studies in lowland South American languages*, 243–271. Austin: University of Texas Press.

Derbyshire, Desmond C. 1985. *Hixkaryana and Linguistic Typology*. (Summer Institute of Linguistics: Publications in Linguistics 76). Dallas: The Summer Institute of Linguistics and the University of Texas at Arlington.

Dickinson, Connie. 2002. Complex predicates in Tsafiki. University of Oregon dissertation.

Dienst, Stefan. 2014. *A Grammar of Kulina*. Berlin: De Gruyter Mouton.

Dixon, R. M. W. & Alexandra Aikhenvald. 1999. Introduction. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The Amazonian Languages*, 1–22. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1986. Noun class and noun classification. In Colette Craig (ed.), *Noun Classes and Categorization,* 105–112. Amsterdam: John Benjamins.

Dixon, R. M. W. 2004. *The Jarawara language of southern Amazonia*. Oxford: Oxford University Press.

Drude, Sebastian. 2014. Reduplication as a tool for morphological and phonological analysis in Awetí. In Gale Goodwin Gómez & Hein van der Voort (eds.), *Reduplication in Indigenous Languages of South America*, 185–215. Leiden: Brill.

Dryer, Matthew. 2007. Noun phrase structure. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume 2: Complex Constructions*, 151–205. 2nd edn. Cambridge: Cambridge University Press.

Duff-Tripp, Martha. 1997. *Gramática del idioma yanesha' (amuesha)*. (Serie Lingüística Peruana 43). Lima: Ministerio de Educación and Instituto Lingüístico de Verano.

Durie, Mark. 1986. The grammaticalization of number as a verbal category. In Vassiliki Nikiforidou, Mary Van Clay, Mary Niepokuj & Deborah Feder (eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, 15–17 February 1986, 355–370. Berkeley: Berkeley Linguistics Society, University of California.

Eberhard, David. 2009. *Mamaindê grammar. A Northern Nambikwara language and its cultural context*. Utrecht: LOT Dissertation Series 236.

Emkow, Carola. 2006. *A grammar of Araona, an Amazonian language of Northwestern Bolivia*. LaTrobe University dissertation.

Epps, Patience. 2008. *A grammar of Hup*. Berlin: Mouton de Gruyter.

Evans, Nicholas. 2006. Dyad constructions. In Keith Brown (ed.), *Encyclopaedia of Language and Linguistics,* 24–28. 2nd edn. Amsterdam: Elsevier Ltd.

Everett, Caleb. 2006. Gestural, perceptual, and conceptual patterns in Karitiâna. Houston: Rice University dissertation.

Everett, Daniel L. & Barbara Kern. 1997. *Wari': the Pacaas Novos language of Western Brazil*. (Descriptive Grammars Series). London & New York: Routledge.

Everett, Daniel L. 1986. Pirahã. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, 200–325. Berlin: Mouton de Gruyter.

Fabre, Alain. 2016. *Gramática de la lengua Nivacle (familia Mataguayo, Chaco Paraguayo)*. München: LINCOM.

Facundes, Sidney da Silva. 2000. *The language of the Apurinã people of Brazil*. Buffalo: SUNY.

Fargetti, Cristina. 2001. *Estudo fonológico e morfossintático da língua juruna*. Universidade Estadual Paulista Julio de Mesquita Filho dissertation.

Fernández Garay, Ana & Graciela Hernández. 2006. *Textos tehuelches: (aonek'o ʔaʔjen): Homenaje a Jorge Suárez*. (Languages of the world: Text collections 24). München: Lincom.

Filho, Sinval Martins de Sousa. 2007. Aspectos Morfossintáticos da Língua Akwẽ-Xerente (Jê). Goiânia: Programa de pos-graduacao em letras e lingüística. Universidade Federal de Goiás dissertation.

Fleck, David. 2003. A grammar of Matsés. Rice University dissertation.

Forster, D. Keith. 2011. *Paya Kuna: An introductory grammar*. SIL International.

Frank, Paul. 1990. *Ika syntax*. Dallas: Summer Institute of Linguistics.

Gabas Jr., Nilson. 1999. A Grammar of Karo, Tupí (Brazil). Ann Arbor: UMI. University of California at Santa Barbara dissertation.

Gakran, Nanblá. 2005. Aspectos Morfossintáticos da Língua Laklãnõ (Xokleng) "Jê". Campinas: Universidade Estadual de Campinas MA thesis.

Galucio, Ana Vilacy. 2001. The Morphosyntax of Mekens (Tupi). University of Chicago dissertation.

Gasparini, Noé. 2011. L'expression du nombre dans les langues tupi-guarani. Université Lumière Lyon 2 MA thesis.

Girón, Jesús Mario. 2008. *Una gramática del wãnsöjöt (puinave)*. Utrecht: LOT Publications.

González, Hebe A. 2005. *A grammar of Tapiete (Tupí-Guaraní)*. University of Pittsburgh dissertation.

Goodwin Gómez, Gale. 1990. The Shiriana dialect of Yanam (northern Brazil). Columbia University dissertation.

Greenberg, Joseph. 1972. Numeral classifiers and substantival number: Problems in the genesis of linguistic type. *Working Papers on Language Universals, No. 9.*

Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In Gunter Senft (ed.), *Systems of nominal classification,* 50–92. Cambridge: Cambridge University Press.

Grondona, Verónica M. 1998. *A Grammar of Mocovi*. Ann Arbor: UMI. University of Pittsburgh dissertation.

Guillaume, Antoine. 2008. *A grammar of Cavineña*. Berlin: Mouton de Gruyter.

Guirardello, Raquel. 1999. *Trumai reference grammar*. Rice University dissertation.

Haas, Mary. 1969. 'Exclusive' and 'inclusive': a look at early usage. *International Journal of American Linguistics* 35(1). 1–6.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2020. *Glottolog 4.3*. Jena: Max Planck Institute for the Science of Human History. (Available online at http://glottolog.org).

Hannβ, Katia & Pieter Muysken. 2014. Reduplication in Andean languages. In Gale Goodwin Gómez & Hein van der Voort (eds.), *Reduplication in Indigenous Languages of South America*, 39–76. Leiden: Brill.

Hardman, Martha J. 1972. Early Use of Inclusive/Exclusive. *International Journal of American Linguistics* 38(2). 145–146.

Hardman, Martha J. 2000. *Jaqaru*. (Languages of the World/Materials 183). München: Lincom.

Hardman, Martha J. 2001. *Aymara*. (LINCOM Studies in Native American Linguistics 35). München: Lincom.

Hart, Helen L. 1988. *Diccionario chayahuita-castellano (Canponanquë nisha nisha nonacaso')*. (Serie Lingüística Peruana 29). Yarinacocha: Ministerio de Educación and Instituto Lingüístico de Verano.

Haspelmath, Martin. 2013. Occurrence of Nominal Plurality. In Matthew Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck

Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/34, Accessed on 2018 09 21.)

Haude, Katharina. 2006. A grammar of Movima. Radboud University Nijmegen dissertation.

Haviland, John B. 1979. Guugu Yimidhirr brother-in-law language. *Language in Society* 8. 365–93.

Henry, Jules. 1948. The Kaingang Language. *International Journal of American Linguistics* 14(3). 194–204.

Ivani, Jessica Katiuscia. 2017. *The morphosyntax of number systems: a cross-linguistic study.* Università degli Studi di Bergamo/Università degli Studi di Pavia Dissertation.

Jensen, Cheryl. 1999. Tupí-Guaraní. In R. M. W. Dixon & A. Y. Aikhenvald (eds.), *The Amazonian Languages*, 125–163. Cambridge University Press.

Jung, Ingrid. 2008. *Gramática del páez o nasa yuwe: Descripción de una lengua indígena de Colombia*. Munich: Lincom Europa.

Klein, Harriet E. Maneslis. 1992. South American languages. In William Bright (ed.), *International Encyclopedia of Linguistics*, Vol. 4, 31–35. Oxford: Oxford University Press.

Krasnoukhova, Olga. 2012. *The noun phrase in the languages of South America*. Utrecht: LOT Dissertation Series 301.

Krasnoukhova, Olga. 2014. Noun Phrase (NP). In Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad. *South American Indigenous Language Structures (SAILS) Online*. Leipzig: Online Publication of the Max Planck Institute for Evolutionary Anthropology. (Available at http://sails.clld.org).

Krasnoukhova, Olga. Forthcoming. Nominal morphological structures and categories of the Central Andes. In Matthias Urban. *The Oxford Guide to the Languages of the Central Andes*. Oxford: Oxford University Press.

Lai, I-Wen. 2009. Time in the Iquito Language. University of Texas at Austin Dissertation.

Lefebvre, Claire M. A. 1975. Person and plural agreement in Cuzco Quechua. Some aspects of variation. University of California at Berkeley dissertation.

Loos, Eugene E. 1999. Pano. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *The Amazonian Languages,* 227–250. Cambridge: Cambridge University Press.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

Manheim, Bruce. 1982. Person, number and inclusivity in two Andean languages. *Acta Linguistica Hafniensia* 17(2). 139–156.

Mattiola, Simone & Spike Gildea. Submitted. *The pluractional marker -pödï of Akawaio (Cariban) and beyond*.

Mattiola, Simone. 2019. *Typology of Pluractional Constructions in the Languages of the World*. (Typological studies in language 125). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Meira, Sérgio. 1999. *A Grammar of Tiriyó*. Houston: Rice University dissertation.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In Hammond, Michael & Michael Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics*, 211–234. San Diego: Academic Press.

Moore, Denny. 1984. *Syntax of the language of the Gavião Indians of Rondônia*. City University of New York dissertation.

Moroz, George. 2017. *Lingtypology: easy mapping for Linguistic Typology*. <URL: https://CRAN.R-project.org/package=lingtypology>.

Mortensen, Charles. 1999. *A Reference Grammar of Northern Embera Languages*. (Studies in the Languages of Colombia 7). SIL International and University of Texas at Arlington.

Mueller, Neele. 2013. *Tense, Aspect, Modality, and Evidentiality Marking in South American Indigenous Languages*. Utrecht: LOT Dissertation Series 324.

Mueller, Neele. 2014. Tense-Aspect-Mood-Evidentiality (TAME). In Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke,

Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad. 2014. *South American Indigenous Language Structures (SAILS) Online*. Leipzig: Online Publication of the Max Planck Institute for Evolutionary Anthropology. (Available at http://sails.clld.org).

Müller, Ana & Luciana Sanchez-Mendez. 2008. Pluractionality in Karitiâna. In Atle Grønn (ed.) *Proceedings of SuB* 12, 442–454. Oslo.

Muysken, Pieter. 1977. *Syntactic developments in the verb phrase of Ecuadorian Quechua*. Dordrecht: Foris.

Muysken, Pieter. 2012. Contacts between indigenous languages in South America. In Campbell, Lyle & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 235–258. Berlin: Mouton de Gruyter.

Muysken, Pieter. Forthcoming. Kallawaya. In Matthias Urban (eds.), *The Oxford Guide to the Languages of the Central Andes*. Oxford: Oxford University Press.

O'Brien, Colleen Alena. 2018. A grammatical description of Kamsá: A language isolate of Colombia. University of Hawaiʻi at Mānoa dissertation.

O'Hagan, Zachary, Lev Michael & Rosa Vallejos. 2013. *Hacia la reconstruccíon morfológica del proto-omagua-kokama*. Paper presented at the Conference on Indigenous Languages of Latin America (CILLA VI) 25 October 2013.

O'Hagan, Zachary. 2011. Proto-Omagua-Kokama: Grammatical sketch and prehistory. University of California at Berkeley MA thesis.

Olawsky, Knut. 2006. *A Grammar of Urarina*. Berlin: Mouton de Gruyter.

Oliveira, Christiane Cunha de. 2005. *The Language of the Apinajé People of Central Brazil*. University of Oregon dissertation.

Overall, Simon. 2017. A *Grammar of Aguaruna (Iiniá Chicham)*. (Mouton Grammar Library, 68). Berlin: De Gruyter Mouton.

Payne, Doris L. & Thomas E. Payne. 1990. Yagua. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, 249–474. Berlin/New York: Mouton de Gruyter.

Payne, Doris. 1987. Noun classification in the western Amazon. *Language Sciences* 9. 21–44.

Payne, Thomas E. & Doris L. Payne. 2013. *A Typological Grammar of Panare: A Cariban Language of Venezuela*. (Brill's Studies in the Indigenous Languages of the Americas). Leiden: Brill.

Peña, Jaime. 2015. A grammar of Wampis. University of Oregon dissertation.

Perri Ferreira, Helder. 2017. Yanomama Clause Structure. Utrecht: LOT Dissertation Series 475.

Quesada, J. Diego. 2007. *The Chibchan languages*. Cartago: Editorial Tecnológica de Costa Rica.

Ribeiro, Eduardo. 2012. A grammar of Karajá. University of Chicago dissertation.

Rich, Rolland. 1975. *Sufijos verbales y apuntes gramaticales sobre el idioma arabela*. (Datos Etno-Lingüísticos 18) Lima: Instituto Lingüístico de Verano.

Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.

Rodrigues, Aryon D. & Ana S. Cabral. 2012. Tupían. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 495–574. Berlin: Mouton de Gruyter.

Rodrigues, Aryon D. 1953. Morfologia do verbo Tupi. *Letras* 1. 121–152.

Romero-Figeroa, Andrés. 1997. *A Reference Grammar of Warao*. (LINCOM Studies in Native American Linguistics 6). München: Lincom.

Rose, Françoise. 2005. Reduplication in Tupí-Guaraní languages: Going into opposite directions. In Bernhard Hurch with editorial assistance of Veronica Mattes (eds.), *Studies on Reduplication*, 351–368. Berlin: Mouton De Gruyter.

Rose, Françoise. 2007. Action repetitive et action répétée: aspect et pluralité verbale dans la reduplication en émérillon. *Faits de Langues* 29. 125–143.

Rose, Françoise. 2011. *Grammaire de l'émérillon Teko, Une Langue Tupi-Guarani de Guyane Française*. Langues et Sociétés d'Amérique Traditionnelle 10. Louvain: Peeters.

Rose, Françoise. 2015. On male and female speech and more: Categorical gender indexicality in indigenous South American languages. *International Journal of American Linguistics* 81(4). 495–537.

Sakel, Jeanette. 2004. *A grammar of Mosetén*. Berlin: Mouton de Gruyter.

Salanova, Andrés. 2014. Reduplication and verbal number in Mẽbengokre. In Gale Goodwin Gómez & Hein van der Voort (eds.), *Reduplication in Indigenous Languages of South America*, 247–272. Leiden: Brill.

Sanches, Mary & Linda Slobin. 1973. Numeral classifiers and plural marking: an implicational universal. *Working Papers on Language Universals* 11. 1–22.

Sanches, Mary. 1971. Numeral classifiers and plural marking: an implicational universal. (Unpublished manuscript).

Sándalo, Filomena. 1997. *A Grammar of Kadiwéu with special reference to the Polysynthesis Parameter*. (MIT Occasional Papers in Linguistics 11). MIT Occasional Papers in Linguistics. Cambridge, Massachusetts: Massachusetts Institute of Technology.

Seifart, Frank & Doris Payne. 2007. Nominal classification in the North West Amazon: Issues in areal diffusion and typological characteristics. *International Journal of American Linguistics* 73. 281–287.

Seifart, Frank & Harald Hammarström. 2017. Language isolates in South America. In Lyle Campbell (ed.), *Language isolates*, 260–286. London & New York: Routledge.

Seifart, Frank. 2005. The structure and use of shape-based noun classes in Miraña (North West Amazon). Nijmegen: Max Planck Institute for Psycholinguistics. Radboud University Nijmegen dissertation.

Seki, Lucy. 2000. *Gramática do Kamaiurá: Língua Tupi-Guarani do Alto Xingu*. Campinas: Editora da Universidade Estadual de Campinas.

Silva, Léia de Jesus. 2011. Morphosyntaxe du Rikbaktsa. Université de Paris 7 – Dénis Diderot dissertation.

Smeets, Ineke. 2008. *A grammar of Mapuche*. (Mouton Grammar Library 41). Berlin: Mouton de Gruyter.

Smith-Stark, Thomas. 1974. The plurality split. *Chicago Linguistic Society* 10. 657–671.

Steele, Susan. 1978. Word order variation: a typological study. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals in Human Language IV: Syntax*, 585–623. Stanford: Stanford University Press.

Stenzel, Kristine. 2013. *A Reference Grammar of Kotiria (Wanano)*. University of Nebraska Press.

Storto, Luciana. 2014. Reduplication in Karitiâna (Tupi). In Gale Goodwin Gómez & Hein van der Voort (eds.), *Reduplication in Indigenous Languages of South America*, 401–426. Leiden: Brill.

Terraza, Jimena. 2009. *Grammaire du Wichí: Morphologie et morphosyntaxe*. Université du Québec à Montréal dissertation.

Thiesen, Wesley & David Weber. 2012. *A Grammar of Bora with Special Attention to Tone*. (SIL International Publications in Linguistics 148). SIL International.

Trillos Amaya, María. 1996. *Categorias grammatical del Ette Taara: lengua de los chimilas*. (Lenguas Aborígenes de Colombia: Descripciones 10). Descripciones. Bogotá: CESO-CCELA, Universidad de los Andes.

Trillos Amaya, María. 1999. *Damana*. (Languages of the World/Materials 207). München: LINCOM.

Urban, Greg. 1985. Ergativity and accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51. 164–187.

Valenzuela, Pilar. 2003. *Transitivity in Shipibo-Konibo Grammar*. University of Oregon dissertation.

Valenzuela, Pilar. 2015. ¿Qué tan "amazónicas" son las lenguas kawapana? Rasgos centro-andinos y evidencia para una posible sub-área lingüística. *Lexis* 39(1). 5–56.

Valenzuela, Pilar. 2018. *Hacia una subárea lingüística intermedia Andes-Amazonia en el piedemonte oriental nor-peruano y sur-ecuatoriano*. Paper presented at Amaznicas-VII, Baños, Ecuador, May 28-June 2, 2018.

Vallejos Yopán, Rosa. 2010. A grammar of Kokama-Kokamilla. Doctoral dissertation, University of Oregon.

Vallejos, Rosa. 2016. *A Grammar of Kukama-Kukamiria: A Language from the Amazon*. Leiden & Boston: BRILL.

van de Kerke, Simon. 2009. Leko. In Mily Crevels & Pieter Muysken (eds.), *Lenguas de Bolivia*. Vol I. *Ambito* Andino, 287–332. La Paz: Plural editores.

van der Voort, Hein. 2004. *A grammar of Kwaza*. Berlin: Mouton de Gruyter.

van der Voort, Hein. 2005. Kwaza in a comparative perspective. *International Journal of American Linguistics* 71. 365–412.

van Gijn, Rik. 2006. A grammar of Yurakaré. Radboud University Nijmegen dissertation.

Van Gysel, Jens E. L. 2017. Temporal predicative particles in Sanapaná and the Enlhet-Enenhlet language family (Paraguay). A Descriptive and Comparative study. Leiden University MA thesis.

Van linden, An. Forthcoming. Harakmbut. In Patience Epps & Lev Michael (eds.), *Amazonian Languages. An International Handbook*. Berlin: de Gruyter Mouton.

Vidal, Alejandra. 2001. Pilagá grammar (Guaykuruan family, Argentina). University of Oregon dissertation.

Vuillermet, Marine. 2012. Grammaire de l'ese ejja, langue tacana d'Amazonie bolivienne. Université Lumière Lyon 2 dissertation.

Weber, David. 1989. *A Grammar Huallaga (Huánuco) Quechua*. (University of California Publications in Linguistics 112). Berkeley and Los Angeles: University of California Press.

Wise, Mary Ruth. 2011. Rastros desconcertantes de contactos entre idiomas y culturas a lo largo de los contrafuertes orientales de los Andes del Perú. In Willem Adelaar, Pilar Valenzuela Bismarck & Roberto Zariquiey Biondi (eds.), *Estudios sobre lenguas andinas y amazónicas: Homenaje a Rodolfo Cerrón-Palomino*, 305–306. Lima: Pontificia Universidad Católica del Perú.

Wojtylak, Katarzyna. 2017. A Grammar of Murui (Bue): A Witotoan language of Northwest Amazonia. James Cook University dissertation.

Xrakovskij, Viktor. 1997. Semanic types of the plurality of situations and their natural classification. In Viktor Xrakovskij (ed.), *Typology of Iterative Constructions*, 3–64. München: Lincom Europa.

Zariquiey Biondi, Roberto. (this volume)

Zariquiey Biondi, Roberto. 2011. *A grammar of Kashibo-Kakataibo*. LaTrobe University dissertation.

Zúñiga, Fernando. 2000. *Mapudungun*. München: LINCOM.

Appendix 1 Sample languages used for the analysis of pronominal number.

Number distinction in personal pronouns: A (=sg/du/pl in all 3 persons); B (=sg/pl in all 3 persons); C (=sg/pl in 1st and 2nd person, but not in 3rd person); D (=sg/pl in 1st person, but not in 2nd or 3rd); E (= no number in personal pronouns)

iso-639-3	Language family	Language name	Number distinction in personal pronouns
arn	Araucanian	Mapudungun	A
apu	Arawakan	Apurinã	B
brg	Arawakan	Baure	B
pab	Arawakan	Parecís	B
tae	Arawakan	Tariana	B
ame	Arawakan	Yanesha'	B
cul	Arawan	Kulina	B
jaa	Arawan	Jamamadí	B
pad	Arawan	Paumari	B
jqr	Aymaran	Jaqaru	B
ayc	Aymaran	Southern Aymara	B
kwi	Barbacoan	Awa-Cuaiquer	B
cof	Barbacoan	Tsafiki	B
boa	Boran	Bora	A
bor	Bororoan	Bororo	B
cbt	Cahuapanan	Shawi	B
hix	Cariban	Hixkaryána	B
pbh	Cariban	Panare	B
tri	Cariban	Trió	B
pav	Chapacuran	Wari	B
cbg	Chibchan	Chimila	A
mbp	Chibchan	Malayo	A
kvn	Chibchan	Border Kuna	B
arh	Chibchan	Arhuaco	B
agr	Chicham	Aguaruna	B
hub	Chicham	Huambisa	B
emp	Chocoan	Northern Emberá	B
ona	Chonan	Selk'nam	B
teh	Chonan	Tehuelche	A
kbc	Guaicuruan	Kadiweu	B
moc	Guaicuruan	Mocoví	B
plg	Guaicuruan	Pilagá	B
amr	Harakmbut	Amarakaeri	C
cht	Hibito-Cholon	Cholón	B
alc	Kawesqar	Qawasqar	E
mtp	Matacoan	Wichí Lhamtés Nocten	B
cag	Matacoan	Nivacle	B
cas	Moseten-Chimane	Mosetén	B

(continued)

iso-639-3	Language family	Language name	Number distinction in personal pronouns
myp	Muran	Pirahã	E
kwa	Naduhup	Dâw	B
jup	Naduhup	Hup	B
wmd	Nambiquaran	Mamaindê	B
sae	Nambiquaran	Sabanê	D
apn	Nuclear-Macro-Je	Apinaje	B
xer	Nuclear-Macro-Je	Akwe-Xerente	B
kpj	Nuclear-Macro-Je	Karaja	B
xok	Nuclear-Macro-Je	Xokleng	B
kgp	Nuclear-Macro-Je	Kaingang	B
rkb	Nuclear-Macro-Je	Rikbaktsa	B
kqq	Nuclear-Macro-Je	Krenak	B
cav	Pano-Tacanan	Cavineña	A
ese	Pano-Tacanan	Ese Ejja	C
aro	Pano-Tacanan	Araona	A
cbr	Pano-Tacanan	Kashibo-Kakataibo	A
mcf	Pano-Tacanan	Matsés	E
sg	Pano-Tacanan	Shipibo-Conibo	B
yaa	Pano-Tacanan	Yaminahua	C
yad	Peba-Yagua	Yagua	A
qub	Quechuan	Huallaga Huánuco Quechua	B
qvi	Quechuan	Imbabura Highland Quichua	B
qus	Quechuan	Santiago del Estero Quichua	B
quh	Quechuan	South Bolivian Quechua	B
cub	Tukanoan	Cubeo	B
gvc	Tukanoan	Kotiria	B
snn	Tukanoan	Siona-Tetete	B
cod	Tupian	Cocama-Cocamilla	B
eme	Tupian	Teko	B
gvo	Tupian	Gavião Do Jiparaná	B
kay	Tupian	Kamayurá	B
ktn	Tupian	Karitiâna	C
arr	Tupian	Karo	B
yrl	Tupian	Nhengatu	B
skf	Tupian	Mekens	B
tpj	Tupian	Tapieté	C
awe	Tupian	Aweti	B
yuq	Tupian	Yuqui	B
cap	Uru-Chipaya	Chipaya	B
huu	Witotoan	Murui	A
wca	Yanomamic	Yanomami	A
shb	Yanomamic	Ninam	A
ayo	Zamucoan	Ayoreo	B
ceg	Zamucoan	Chamacoco	B
arl	Zaparoan	Arabela	A
iqu	Zaparoan	Iquito	B

kbh	isolate	Camsá	A
con	isolate	Cofán	B
ito	isolate	Itonama	B
kxo	isolate	Kanoê	B
xwa	isolate	Kwaza	C
lec	isolate	Leko	B
omc	isolate	Mochica	B
mzp	isolate	Movima	B
myr	isolate	Muniche	C
pbb	isolate	Páez	B
pue	isolate	Puelche	A
pui	isolate	Puinave	B
tpy	isolate	Trumai	A
ura	isolate	Urarina	B
auc	isolate	Waorani	A
yuz	isolate	Yuracaré	B

Roberto Zariquiey

18 The category of number in Kakataibo (Panoan)

Abstract: In this chapter I discuss number in Kakataibo, a Panoan language spoken by approximately 3000–3500 people in the Peruvian Regions of Huánuco and Ucayali. Number is manifested in Kakataibo in the nominal domain, the verbal domain and in other lexical categories (particularly, numerals and quantifiers). In addition to a general plural marker, there are three bound nominal morphemes that express number-related categories: a distributive marker, a generic marker, and a collective marker. Interestingly, personal pronouns have more number choices than nouns, including a distinction between dual/paucal and plural, although this distinction is no longer made by younger speakers. In the verbal domain there is a plural marker, -*kan*, as well as various other number-related verbal markers: a distributive affix, a marker for plurality of objects, two 'adverbial' affixes meaning 'once' and 'again', and two iterative markers, which also express different types of spatial trajectories. Iteration may also be expressed through verbal reduplication of various types (reduplication of the root, reduplication of the stem and reduplication of the word). Kakataibo has small inventories of numerals and quantifiers. Some number-related categories do not have dedicated markers or constructions in Kakataibo: for example, the language lacks dedicated constructions for associative plurals and dyadic kinship terms.

1 Overview

Kakataibo is a Panoan language spoken by approximately 3000–3500 people in the Peruvian Regions of Huánuco and Ucayali (see Figure 1). Kakataibo is the westernmost Panoan language and therefore the one closest to the Andes Mountains. Various authors coincide in treating this language as the only member of its branch (Shell 1965, D'Ans 1973, Loos 1999, Fleck 2013). As discussed in Zariquiey (2011a), there are four extant Kakataibo dialects, spoken along the lower Aguaytía, upper Aguaytía, Sungaroyacu and San Alejandro Rivers. 'Nokamán', a variety named and minimally documented by Tessmann (1930), was a fifth dialect of Kakataibo, now extinct (Zariquiey 2013). The data presented in this paper come from the Lower Aguaytía Variety (as described in Zariquiey 2011b and 2018).[1]

---

1 The examples in this paper are given in the Kakataibo practical orthography: (IPA symbols are enclosed in square brackets where these differ from the practical representation): a, e, ë [ɨ], i, o, u, p, t, k, ku [kʷ], b [β], r [ɾ], m, n, ñ [ɲ], s, sh [ʃ], x [ʂ], ts and ch [tʃ]. I also use <'> to represent the contrastive glottal stop [ʔ], which is found at the beginning of certain words.
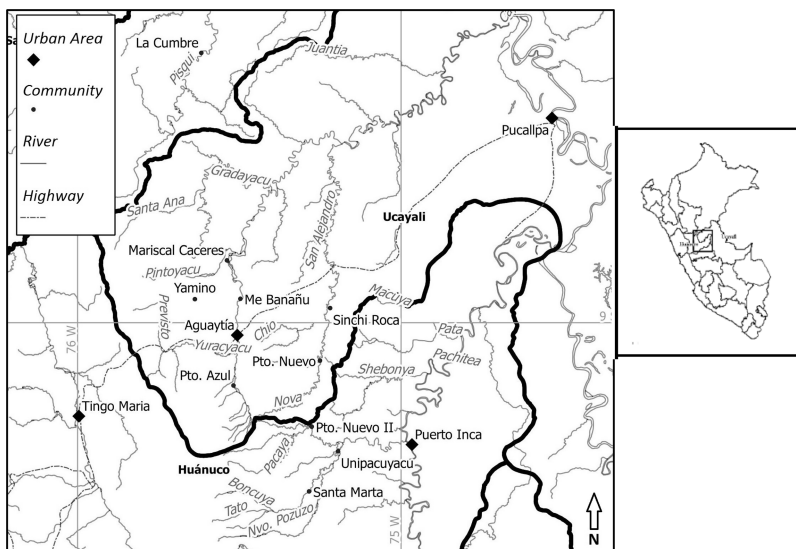
Fig. 1: Location of major Kakataibo settlements.

Kakataibo is a predominantly postpositional language in which almost all the grammatical categories are expressed by suffixes, enclitics or postpositions. The only exception is a closed set of prefixes, mainly related to body parts (see Zariquiey and Fleck 2012).

Kakataibo is predominantly agglutinative and morpheme boundaries are generally transparent. See the example in (1), where the verb root *ux* 'sleep' is followed by five bound morphemes, including the plural marker *-kan*, which will be discussed in detail in this chapter. Fusional affixes are by far less common in Kakataibo. One example of a *portmanteau* verbal morpheme is the suffix *-kian*, which is a single morphological unit that encodes tense, aspect and subject cross-reference, as illustrated in (2).

(1) *ux-**mi-tëkën-kan-i-n***
sleep-**CAUS-again-PL-IPFV-1/2**
'We are making (someone) sleep again.'

(2) *ñui-xun-**kian***
say-**BEN-REM.PST:HAB:3**
'(S)he used to tell tales to somebody.'

As becomes clear from the examples in (1) and in (2), words in Kakataibo (particularly verbs) may be produced by the combination of two or more morphemes, so the language exhibits some tendency towards synthetic structures (see Zariquiey 2018:

Section 4.2.2 for a discussion of morphological complexity in Kakataibo). However, imperative verbal forms, S/P (i.e. absolutive) nouns and P pronouns remain unmarked. In (3), which illustrates an imperative construction, the pronoun *a* '3SG', the noun *kuriki* 'money' and the verb *'inan* 'give' remain unmarked.

(3) ***a*** *ka* **kuriki** ***'inan***
    3SG:P NAR **money:ABS give:IMP**
    'Give him money!'

The language has both head and dependent marking, and a complex system of grammatical relations. Case marking on pronouns and topical noun phrases (henceforth NPs) follows a tripartite alignment (S ≠ P ≠ A), non-topical NPs exhibit ergative case marking (S = P ≠ A), second and third person emphatic pronouns are always unmarked (S = P = A) and first person emphatic pronouns present a horizontal alignment type (S ≠ P = A). Argument indexation on verbs and by second-position enclitics, in turn, exhibits an accusative alignment type (S = A ≠ P). The tripartite case alignment associated with pronouns (=*n* 'A', =*x* 'S' and unmarked 'P') and the ergative alignment attested on NPs (=*n* 'ergative' and unmarked absolutive) are illustrated in the examples in (4) and (5) (argument indexation on verbs and second-position enclitics is also illustrated in these examples). All case markers in Kakataibo are enclitics that operate over the NP, rather than over independent words. The ergative case marker exhibits a fairly complex allomorphy (in (5a) it surfaces as =*nën*). For a complete discussion of grammatical relations in Kakataibo, see Zariquiey (2018: Section 15.2).

(4) a. *mi=**n*** *kamina* ***'ë*** *mëë-a-n*
       2SG=**A** NAR:2 1SG:**P** hit-PFV-1/2
       'You hit me.'

    b. *mi=**x*** *kamina* *'ux-a-n*
       2SG=**S** NAR:2 sleep-PFV-1/2
       'You slept.'

(5) a. *Ricardo=**nën*** *ka* *chuna* *'a-a-x-a*
       Ricardo=**ERG** NAR:3 spider.monkey:ABS kill-PFV-3-NON.PROX
       'Ricardo killed a/the spider monkey.'

    b. ***Ricardo*** *ka* *'ux-a-x-a*
       Ricardo:ABS NAR:3 sleep-PFV-3-NON.PROX
       'Ricardo slept.'

As in other Panoan languages (see Loos 1999), verbs in Kakataibo are lexically transitive or intransitive, with only 2 ambitransitive verbs encountered so far in the

language. Verbs exhibit interesting processes of transitivity harmony and transitivity-sensitive affixation. Such processes constitute one of the core properties of the syntax of Kakataibo (and Panoan, in general).

Word order is pragmatically-oriented, but there is a tendency for sentences to be verb-final. There is no fixed order in the noun phrase either, and most modifiers (including numerals and quantifiers) are allowed to appear after or before the head. Other relevant grammatical features of Kakataibo are the existence of a rich system of switch-reference used in clause-chaining (see Zariquiey 2016) and the pervasive use of grammatical nominalizations in discourse (see Valle and Zariquiey 2019).

Number is manifested in Kakataibo in the nominal domain, the verbal domain and in other lexical categories (particularly, numerals and quantifiers). In the nominal domain there are four bound morphemes that express number-related categories: the plural marker =*kama*, the distributive marker =*ti(i)bi*, the generic marker -*ina(k)* and the collective marker -*baë*. Personal pronouns have more number choices than nouns, including a distinction between dual/paucal and plural, but this distinction is no longer made by younger speakers. In the verbal domain there is a plural marker, -*kan*, as well as various other number-related verbal markers that are also discussed in this chapter: -*rabat* 'distributive', -*pat* 'plural objects', -*tabat* 'once', -*tëkën* 'again', and two iterative markers, -*bëkin* and -*uku*, which also express different types of spatial trajectories. Iteration may also be expressed through verbal reduplication of various types (reduplication of the root, reduplication of the stem and reduplication of the word). Kakataibo has small inventories of numerals and quantifiers. Some number-related categories do not have dedicated markers or constructions in Kakataibo: for example, the language lacks dedicated constructions for associative plurals and dyadic kinship terms.

The present chapter is organized as follows. In Section 2, I present information on the manifestations of the category of number in Kakataibo, including pronouns (Section 2.1), nouns (Section 2.2, which includes references to numerals and quantifiers) and verbs (Section 2.3, which includes discussion on lexical plurals, morphological number markers and reduplication). Section 3 presents some brief notes on the syntax of number in Kakataibo, Section 4 comments on the discursive properties of number in Kakataibo, and Section 5 presents the conclusions of this study.

2 Pronominal, nominal and verbal number

2.1 Pronominal number

As discussed in Zariquiey (2018: 175–177), Kakataibo has two different non-singular forms for each pronoun. Older speakers attribute specific number meanings to each pronominal form, resulting in a complex system with three categories: singular, dual/paucal and plural (plus an exclusive/inclusive distinction in the first-person

plural). Younger speakers are familiar with the different non-singular forms attested for each pronoun, but treat them equally as non-singular forms, without establishing further number or clusivity distinctions.

The existence of two non-singular forms for each pronoun seems to be related to two different systems for marking number that co-exist in the language: one that is reminiscent of a more archaic system reconstructed in Zariquiey (2006) for proto-Panoan (with different plural markers for each person and a clusivity distinction for the first person plural); and another based on the combination of the general Kakataibo plural marker *=kama* with all the pronouns in the language. Elderly Kakataibo speakers systematically interpret the pronominal forms with the archaic markers as dual (in the case of the first and second person) and paucal (in the case of third person forms). The plural pronouns formed with the general plural marker *=kama* are interpreted by elderly speakers as referring to larger groups. In discourse, however, archaic forms may be used for plural referents, and pronouns carrying the plural marker *=kama* can also be used for dual/paucal referents as well, making the distinction between these number categories less clear-cut. In turn, young Kakataibo speakers are not familiar with such number distinctions and use pronominal forms marked with the archaic plural markers and those marked with the general plural maker *=kama* as synonyms.

The clusivity distinction attested among elderly Kakataibo speakers relates to the use of the specialized first person plural pronoun *nu*, which is interpreted as inclusive, and the possibility of adding the plural marker *=kama* to the singular form of the first person pronoun *'ë*. The resulting form, *'ëkama,* is interpretable only as an exclusive plural pronoun, whereas *nu(=kama)* is an inclusive pronoun. Again, this distinction is not clear-cut, since the dedicated first-person plural pronoun *nu* is used for both inclusive and exclusive referents. The pronominal form *'ë=kama* 'first person plural exclusive' was identified only by elderly speakers in elicitation and it became clear that they would never use this form for inclusive reference under any circumstance. Young speakers considered this form pragmatically marked or even ungrammatical. Table 1 compares the pronominal systems attested among elderly and young Kakataibo speakers, respectively. The differences between the two paradigms suggest that the Kakataibo pronominal system is undergoing a rapid process of paradigmatic change.

Reflexive pronouns, as well as interrogative pronouns and demonstratives, exhibit only two number categories: singular and plural (plural on these word classes is encoded by the general plural marker *=kama*). Other types of pronouns (negative and indefinite) are invariantly singular. As an illustration of these patterns, the examples in (6) and (7) feature the interrogative pronouns *ui* 'who' and *a ñu* 'what' in their plural forms. In turn, the examples in (8a) and (8b) feature the singular and the ungrammatical plural form of the negative pronoun *uibi* 'nobody', respectively.[2]

---

2 Note that negative pronouns are obtained in Kakataibo by combining interrogative pro-forms like *ui* 'who' with the emphatic clitic *=bi*.

Tab. 1: Kakataibo personal pronouns: elderly vs. young speakers.

Person	Elderly speakers	Young speakers
1 singular	*'ë*	*'ë*
1 dual inclusive	*nu*	*nu ~ nukama*
1 plural inclusive	*nukama*	
1 plural exclusive	*'ëkama*	
2 singular	*mi*	*mi*
2 dual	*mitsu*	*mitsu ~ mikama*
2 plural	*mikama*	
3 singular	*a*	*a*
3 paucal	*atu*	*atu ~ akama*
3 plural	*akama*	

As indicated in the example, the plural negative form is ungrammatical (an identical situation will be found for indefinite pronouns).

(6) **ui=kama=n** *kara*   *Roberto*   *pi-mi-a-x-a?*
  who=PL=A NAR:INT:3 Roberto:ABS eat-CAUS-PFV-3-NON.PROX
  'Who (plural) fed Roberto?'

(7) **a_ñu=kama** *kara*   *Ricardo=nën 'imainun Wilton=nën*
  what=PL   NAR:INT:3 Ricardo=ERG and   Wilton=ERG
  pi-mi-a-x-ín
  eat-CAUS-PFV-3-PROX
  'What (plural) did Ricardo and Wilton feed (him)?'

(8) a. **uibi**   *kana 'ë=n   'is-a-n*
    nobody NAR:1 1SG=A see-PFV=1/2
    'I didn't see anybody.'

   b. ***uikamabi** *kana 'ë=n   'is-a-n*
    nobody:PL NAR:1 1SG=A see-PFV=1/2
    'I didn't see anybody (plural).'

2.2 Nominal number

2.2.1 Count vs. non-count nouns

Most nouns in Kakataibo have access to pluralization by means of the general plural marker of the language, *=kama*, discussed in Section 2.2.2.1. There is, however, a short list of nouns (only four to my knowledge) that cannot be morphologically pluralized by any means (or modified by a numeral or quantifier), and therefore can be analyzed as non-count nouns. The four non-count nouns attested in Kakataibo are listed in (9).

(9) Nouns that cannot be pluralized or combined with a numeral (*rabĕ* 'two')
    uñe 'rain'   **uñe=kama*   *rabĕ uñë
    kuin 'cloud' **kuin=kama* *rabĕ kuin
    masi 'sand'   **masi=kama* *rabĕ masi
    ni 'jungle'   **ni=kama*   *rabĕ ni

The distinction between count and non-count nouns has been reported for Shipibo-Konibo, another Panoan language, by Valenzuela (2003b: 204), who states that: "[n]on-count nouns do not combine with numerals and cannot take the plural *-bu*". Valenzuela includes in this noun class forms like *tashi* 'salt', *jënë* 'flowing water', *unpax* 'still water', *ui* 'rain', *wakanawa* 'school of fish', *kuin* 'cloud(s)', *niwë* 'wind', *bëchun* 'wave', *mashi* 'sand', *mai* 'land', *manu* 'mud', *nai* 'sky', *manish* 'weed'. In addition, food and drink products like *arus* 'rice,' *atsa putu* 'manioc flour', *bata* 'sugar' and *bëxnan* 'sugar cane liquor' are non-count nouns in Shipibo-Konibo, and this is also the case of some body-part nouns associated with hair and body-hair, such as *rani* 'body hair,' *bëru karani* 'eyebrow,' *bëru këxni* 'eyelash', *këni* 'beard, moustache' and *buu* 'hair' (see Valenzuela 2003b: 204).

Most of the words mentioned by Valenzuela have cognate forms in Kakataibo, but they can be counted and pluralized. For example, the word *me* 'earth' (*mai* in Shipibo-Konibo) is used in both plural and singular forms, as shown in the text example in (10).

(10) *usa*   *'ain* *ka*   *nukama presidente Prado a=n* *ka*   *nu*
     like.that being NAR:3   1PL   president Prado 3SG=A NAR:3 1PL
     'inan-a-x-a   *me* *cha=ira* **me=kama** *'inan-akë-x-a*
     give-PERF-3-NON.PROX land big-INTF   **land=PL**   give-REM.PST-3-NON.PROX
     'Then, the president Prado gave us a big land, several lands.'

The existence of fewer non-count nouns in Kakataibo than in Shipibo-Konibo requires further explanation. One fact that may help to understand why the class of non-count nouns are radically different in terms of their size in the two languages

may have to do with the fact that Shipibo-Konibo and Kakataibo exhibit different plural markers with different diachronic stories. Indeed, the Shipibo-Konibo plural marker =*bo* is diachronically related to the Kakataibo generic marker -*baë*, which is a collective marker with a more restricted distribution. Understanding the history of this marker, as well as the history of the general Kakataibo plural =*kama* (likely cognate with the verbal plural -*kan*, which is found in most Panoan languages) may contribute to an overall understanding of the category of number in the family.

2.2.2 Number markers

In what follows, I discuss and illustrate the four number markers that operate over nouns and NPs. As can be seen in Table 2, the list comprises two NP enclitics (=*kama* 'general plural' and =*ti(i)bi* 'distributive') and two derivational suffixes (-*ina(k)* 'generic plural' and -*baë* 'collective').

Tab. 2: Kakataibo nominal number markers.

Marker	Meaning	Morphological nature
=*kama*	'general plural'	NP enclitic
=*ti(i)bi*	'distributive'	NP enclitic
-*ina(k)*	'generic plural'	Derivational suffix
-*baë*	'collective'	Derivational suffix

2.2.2.1 The general plural *=kama*

The marker =*kama* is optional and is used only if the speaker wants to be explicit about the number of an NP. It is common to find NPs without =*kama* when talking about plural referents if the number can be inferred from the context. On the other hand, NPs marked by =*kama* cannot have singular interpretations under any circumstance. This is exemplified in (11) and (12).

(11) *xanu=**kama*** *ka* *kwan-akë-x-a*
    woman=**PL**:ABS NAR:3 go-REM.PST-3-NON.PROX
    'The women went.' (*'the woman went')

(12) *xanu* *ka* *kwan-akë-x-a*
    woman NAR:3 go-REM.PST-3-NON.PROX
    'The woman went / The women went.'

Even though =*kama* 'plural' is not obligatory, it is very frequent in natural texts. As illustrated in the text example in (13), =*kama* is a NP enclitic in the sense that it appears at the right edge of NPs (as in [*nun aintsi bërama*]=*kama* 'our ancestor=PL').

(13) *kananuna ësa-o-kin*   *bana*   *ñui-xun-kin*   [*nun*
     NAR:1PL   like.this-FACT-S/A>A:SE   tale:ABS   tell-BEN-HAB.PST:1/2   we=GEN
     aintsi   *bërama*]=***kama*** *'i-a*
     relative   old**=PL**:ABS   be-NMLZ
     'Doing like this, we used to tell the stories about what our old relatives were.'

Proper names cannot combine with =*kama*. Any combination of this sort is ungrammatical in the language.

2.2.2.2 =*ti(i)bi* 'distributive'

The enclitic =*ti(i)bi* 'distributive' indicates that the event is associated independently with each of the individuals referred to by the NP. It is rare in natural texts, but examples with this enclitic are readily provided by speakers in elicitation. It can occur with or without the plural enclitic =*kama*, as illustrated in (14) and (15). Note that =*ti(i)bi* follows the plural maker (14), as well as case enclitics (15).

(14) *'ë=x*   *kana*   *xanu=bë=**tibi***   *Lima=nu kwan-ti 'ain*
     1SG=S   NAR:1SG   woman=COM:S=**DIST**   Lima=DIR go-NMLZ be:1/2
     'I will go to Lima with each woman.'

(15) *xanu=kama=**tiibi** ka*   *Lima=nu kwan-ti 'ikën*
     woman=PL**=DIST** NAR:3   Lima=DIR go-NMLZ be:3
     'Each woman will go to Lima.'

2.2.2.3 -*ina(k)* 'plural generic'

The meaning of -*ina(k)* is not easy to determine and further research is required to circumscribe its semantics more precisely. Its meaning seems to revolve around two basic components: the referents are expected to be both (i) generic and (ii) plural (Zariquiey 2018: 227). Thus, nouns modified by -*ina(k)* refer to a whole class of individuals (as generic plurals in English in examples like 'dogs are good friends') and are inherently plural, in the sense that forms modified by -*ina(k)* cannot be modified by numerals like *achushi* 'one' or quantifiers like *itsama* 'few' (with some exceptions commented below). Examples of the use of this suffix are provided in (16).

(16) *anu-xun*   *nukën*   *bakë_bëchikë nukën*   *ini_bëchikë*
     there-S/A>A:SE 1PL:GEN son:ABS   1PL:GEN daughter:ABS
     *'ó-**ina***   *chuna-**ina***   *pi-tankëxun kaniot-i-n*
     tapir-**GENER:ABS** monkey.sp-**GENER:ABS** eat-S/A>A:PE raise-IPFV-1/2
     'There, we will raise our sons and daughters, eating tapir and monkey.'

The suffix -ina(k) is mostly used in combination with nouns that denote animal species, but it can also be used on three nouns that do not refer to animals when used without the suffix: *ñu* 'thing', *me* 'earth' and *baka* 'river'. In combination with this suffix, these nouns refer to generic ethnobiological taxa that cover various species: *ñuina* 'edible animals in general', *mena* 'animals which make holes in the ground and live there' and *bakaina* 'all the kinds of aquatic animals'. The three forms just listed can be modified by numerals and quantifiers. The suffix -ina(k) cannot appear in combination with the plural marker =*kama*, and it always appears before case enclitics, as is the case of -*baë* 'collective' in (17). Etymologically, the suffix -ina(k) 'generic' comes from the lexical form **inaka* which meant something like 'domesticated (animal)' in proto-Panoan.

2.2.2.4 -*baë* 'collective'

The distribution of -*baë* 'collective' is restricted to (dead) ancestors and to generic animal species modified by -*ina(k)* 'plural generic' (see Section 2.3.2.3). Any attempt to modify any other noun with this collective marker will result in an unacceptable form. Some examples of the suffix -*baë* are presented in (17).

(17) *usa-i*   *ka*   *ëo kamë-o-akë-x-a*   *nu=n*
     like.that-S/A>S:SE NAR:3 a.lot.of-FACT-REM.PST-3-NON.PROX 1PL=GEN
     '*anibu*   *nukën*   *chaiti-baë=n*   *nukën*   *xuta-baë=n*
     ancestor 1PL:GEN ancestor-**COL**=ERG 1PL:GEN grandfather-**COL**=ERG
     'Then, our ancestors reproduced themselves a long time ago.'

This suffix shows some semantic overlap with the inflectional plural marker =*kama* (presented in Section 2.3.2.1), but the plural marker shows a much wider distribution, and can appear with almost all nouns in the language (e.g., *xubu=kama* 'houses' but not **xubu-baë*). The plural marker =*kama* can modify kinship terms too: thus, we find, for example, both *chichibaë* '(dead) grandmothers' and *chichikama* 'grandmothers' (more on this is presented in Section 4).

2.2.3 On the distributional properties of nominal number markers in Kakataibo

Kakataibo NPs are headed by a noun, which may or may not be accompanied by one or two modifiers (NPs with more than two modifiers are rare in natural speech but readily produced in elicitation). NP-modifiers in Kakataibo include demonstratives, adjectives, numerals, quantifiers, bare nouns and genitive phrases. All NP-modifiers can appear in the pre-head position. Demonstratives, adjectives, numerals and quantifiers can also appear in the post-head position. Genitives and modifying nouns cannot appear after the nominal head, under any circumstance. All this is summarized in Table 3.

Tab. 3: Distributional possibilities of NP-modifiers.

Type of modifier	pre-head	post-head
demonstratives	YES	YES
adjectives	YES	YES
numerals/quantifiers	YES	YES
bare nouns	YES	NO
genitive phrases	YES	NO

The suffixes *-ina(k)* 'generic' and *-baë* 'collective' always appear on nouns, irrespectively of their position within the NP. A completely different situation is found, however, regarding the enclitics *=kama* 'plural' and *=ti(i)bi* 'distributive', which attach to the rightmost element of the NP, either the head or any type of post-head modifier. That is, on their post-head position, NP-modifiers may host either of the two number markers analyzed as clitics in this chapter. This is illustrated with a demonstrative and an adjective in the examples in (18) and (19).

(18) [*tsatsa ënë*]*=tiibi*   *ka* *pi'!*
     fish.sp this=**DIST**:ABS IMP eat:IMP
     'Eat each of these fishes!'

(19) [*ñu*   *upí*]*=kama*   *kaina kuëën-i-n?*
     fish.sp beautiful=**PL**:ABS INT:2 want-IPFV-1/2
     'Do you want the beautiful things?'

Under the necessary pragmatic conditions, demonstratives, some numerals, quantifiers and adjectives can head NPs. When this happens, they may also attach the number enclitics directly, as in the examples (20) and (21).

(20) *ënë=tiibi*   *ka* *pi'!*
     this= **DIST**:ABS IMP eat:IMP
     'Eat each of these (ones)!'

(21) *upí=kama*   *kaina kuëën-i-n?*
     beautiful=**PL**:ABS INT:2 want-IPFV-1/2
     'Do you want the beautiful (ones)?'

Thus, word classes other than nouns can carry the number enclitics, but only when they are post-head NP-modifiers or NP-heads. Note that *=kama* 'plural' is more widespread than *=ti(i)bi* 'distributive'.

2.2.4 Some notes on numerals and quantifiers

The numerical system of Kakataibo is very simple, as has been reported for Amazonian languages in general (Dixon and Aikhenvald 1999). There are just three numerals in the language: *achushi* 'one', *rabë̆* 'two' and *mapai* 'three' or 'five' (depending on the speaker). The meaning 'five' can alternatively be expressed with *mëkën* 'hand'. In fact, the word *mapai* is not known by younger Kakataibo speakers. Other small numbers can be expressed by combining the three forms presented above (often by adding the conjunction *'imainun* 'and').

Kakataibo has only two words that can clearly be identified as quantifiers: *'itsa* 'a lot' and *kamabi* 'all'. The form *kamabi* 'all' can be analyzed as the combination of the plural marker *=kama* and the emphatic clitic *=bi,* but it is a fully independent word. The form *'itsa* 'a lot' can also be used in its negative form to mean 'not many/ much, a few' (i.e., *'itsa=ma=shi* 'a lot=NEG=only'). Quantifiers, as numerals, cannot be used with the few uncountable nouns attested in Kakataibo.

Numerals and quantifiers can be used as NP modifiers or as NP heads (when the context allows for the proper interpretation). The numeral *achushi* 'one' functions as an indefinite marker, in a fashion similar to what is found in languages like Spanish. The following examples provide some illustration of the use of numerals and quantifiers in Kakataibo. In (22) the quantifier *'itsa* 'a lot' is the head of a constituent equivalent to an NP, whereas in (23) the numeral *achushi* is used as the modifier of the NP head *uni* 'man'. It is important to note that the plural marker *=kama* is not allowed in NPs including a numeral (other South American languages with similar patterns in which numerals ban number marking include Cuzco Quechua and Mosetén; Krasnoukhova 2012: 104 ff; see also Krasnoukhova, this volume).

(22) **'itsa**   *kana*   *is-a-n*
     a.lot:ABS NAR:1SG see-PFV-1/2
     'I saw a lot.'

(23) **achushi** *uni=n*   *ka*   *'a-a-x-a*
     one   man=ERG NAR:3 do-PFV-3-NON.PROX
     'One/a man did (it).'

Numerals and quantifiers can be reduplicated, as an intensification strategy. This is shown in (24), where the form *'itsa* 'a lot' is reduplicated.

(24) *y*   *usa-o-kin*   *ka*   *'a-akë-x-a*   **'itsa 'itsa**
     and like.that-FACT-S/A>A:SE NAR:3 do-REM.PST-3-NON.PROX **a.lot a.lot**
     forma
     form:ABS
     'And doing like that, they prepared lots of forms (of arrows) a long time ago.'

2.3 Verbal number

Number in the Kakataibo verb is expressed through (at least) three different morphological strategies. There are a few suppletive verb pairs based on number (singular S vs. plural S), there are various suffixes that relate to number, and there is productive verbal reduplication. These three strategies are discussed in the following subsections.

2.3.1 Lexically plural verbs

A few predicates exhibit pairs of verbs that are distinguished by number in Kakataibo. This is the case of the pairs *kuan* 'go (singular)' and *ri* 'go (plural)', and *u* 'come (singular) and *bëkan* 'come (plural)' (note that *bëkan* has lexicalized the plural marker -*kan*, discussed in section 2.3.2.1.1). Only *ri* 'go' can be combined with the plural marker -*kan*, but there is no obvious semantic difference in association with the presence/absence of the suffix. The meaning of *ri* 'go (plural)' and *bëkan* 'come (plural)' is collective rather than distributive. Similar verb pairs are also found in other Panoan languages (see Valenzuela 2017 for more details). To my knowledge the pairs just listed are the only ones attested in Kakataibo, a language in which lexically plural verb roots are not as common as in other Panoan languages. NPs acting as the subject of lexically plural verbs may or may not carry a plural marker.

2.3.2 Morphological number in the verb

In addition to the verbal plural marker -*kan*, there are six other verbal suffixes that can be analyzed as related to number. As discussed below, some relate to number in association with (one of) the arguments and the others express number in relation to the event. The morphological number markers that occur on the verb are summarized in Table 4.

Tab. 4: Kakataibo verbal number markers.

Marker	Meaning	Semantic type
-*kan*	'plural'	Argument number
-*rabat* ~ -*abat*	'distributive'	Argument number
-*pat*	'plural objects'	Argument number
-*taba*	'once'	Event number
-*tëkën*	'again'	Event number
-*bëkin*	'iterative in different places'	Event number
-*uku*	'iterative in one direction'	Event number

2.3.2.1 Number in association with the arguments

2.3.2.1.1 *-kan* 'plural'

The plural marker *-kan* is used only for third person subjects, so plurality of first and second person subjects is only specified in the form of the independent pronoun. As is shown by the comparison of the examples in (25a–b), which feature a transitive verb, and (26a–b), which illustrate the use of the suffix on an intransitive predicate, *-kan* 'plural' is not obligatory and plurality of third person subjects does not need to be indicated on the verb. The plural marker on NPs *=kama* and the plural marker on verbs *-kan* are likely to be diachronically related: the verbal plural marker may be a phonologically reduced and more bound version of the nominal one (it is a suffix and not an enclitic): *=kama > -kan*.[3]

(25) a. *nukën   chaitiokë=kama=n kaisa   'ó   nami*
     1PL:GEN ancestor=PL=ERG   NAR:REP:3 tapir meat:ABS
     *pi-**kan**-akë-x-a*
     eat-**PL**-REM.PST-3-NON.PROX
     'It is said that our ancestors ate tapir meat a long time ago.'

   b. *nukën   chaitiokë=**kama**=n kaisa   'ó   nami*
     1PL:GEN ancestor=**PL**=ERG   NAR:REP:3 tapir meat:ABS
     pi-akë-x-a
     eat-REM.PST-3-NON.PROX
     'It is said that our ancestor(s) ate tapir meat a long time ago.'

(26) a. *nukën   chaitiokë=kama kaisa   ënë baka=nu*
     1PL:GEN ancestor=PL   NAR:REP:3 this river=LOC
     *tsó-**kan**-akë-x-a*
     live-**PL**-REM.PST-3-NON.PROX
     'It is said that our ancestors lived in this river at a long time ago.'

   b. *nukën   chaitiokë=kama kaisa   ënë baka=nu*
     1PL:GEN ancestor=PL   NAR:REP:3 this river=LOC
     tsó-akë-x-a
     live-REM.PST-3-NON.PROX
     'It is said that our ancestors lived in this river at a long time ago.'

---

3 As indirect evidence in favor of this analysis, it is important to note that a similar morphophonological pattern (mV# ~ n#) is also attested in relation to some body-part nouns. Body-part nouns in Kakataibo exhibit corresponding shortened forms, which are called 'body-part prefixes' in the Panoan literature (see Zariquiey and Fleck 2012). Some so-called body-part prefixes exhibit the alternation pattern mV# / n# when contrasted with their corresponding full noun, as in *xama* / *xan-* 'new unopened (palm) frond' and *tamu* / *tan-* 'cheek'.

2.3.2.1.2 *-(r)abat* 'distributive'

The suffix *-(r)abat* 'distributive' is inherently plural, but differently from the plural marker *-kan*, *-(r)abat* indicates that the action is being carried out by various individuals independently. The primary meaning of this suffix is thus a strong individuation of every member in the group that carries out the event, such that the event is conceptualized as being composed of various individualized actions. One example of *-(r)abat* 'distributive' is presented in (30).

(27) *bari-i*   *rit-kian-pun-i-a*
     look.for-S/A>S:SE go.together-going.INTR-PST.HOURS-IPFV-NON.PROX
     kaisa   *kain-xun*   *kain-këx=bi*   *kaisa*
     NAR:REP:3 wait-S/A>A:SE wait-P>S:PE=same NAR:REP:3
     *sharat-**rabat**-i*   *ri-kwatsin-akë-x-ín*
     make.noise-**DIST**-S/A>S:SE go.together-coming:INTR-REM.PST-3-PROX
     'It is said that, when he was waiting for (his enemies), the ones who went
     early that day to look for animals came together, each of them making noise.'

2.3.2.1.3 *-pat* 'plural objects'

The Kakataibo directional verbal suffix *-pat* 'downward, transitive' can be used to indicate that the P of a transitive event is plural (this suffix is similar to the semantically equivalent morpheme *-pake* in Shipibo-Konibo, which is described as having a distributive function similar to the one described here; Valenzuela 2010). The use of the marker *-pat* indicating 'plural objects' is illustrated in (28).

(28) *ain xëni 'a-ti*   *ka raëska-**pat***
     their fat   cook-NMLZ IMP singe-**PL:P**
     'Singe (all these animals) in order to cook their fat!'

2.3.2.2 Number in association with the event

2.3.2.2.1 *-taba* 'once'

The suffix *-taba* 'once' is used to indicate that the event was conducted only one time or for the first time, as illustrated in (29).

(29) *'ë=n*   *kana*   *chaxu nami*   *pi-**taba**-t-a-n*
     1SG=ERG NAR:1SG deer   meat:ABS eat-**once**-HAR-PFV-1/2
     'I once ate deer meat.'

2.3.2.2.2 *-tëkën* 'again'

The suffix *-tëkën* 'again' is frequently used in texts and natural speech to present the second occurrence of an event, presupposing that the same event happened

earlier for the first time. In the example in (30), the meaning 'again' is clear, since the first occurrence of the event is presented as a subordinate predicate.

(30) *'a-mi-kin*   *'a-mi-pun-kin*   *kaisa*
do-CAUS-S/A>A:SE do-CAUS-PST.HOURS-S/A>A:SE NAR:REP:3
ñantan-bu-këbëtan   *'a-mi-**tëkën**-akë-x-ín*
get.dark-advanced-DS/A/P:SE:TRAN do-CAUS**-again**-REM.PST-3-PROX
'It is said that, having made (them) do it earlier, when it got dark, they made them do it again.'

2.3.2.2.3 *-bëkin* 'iterative in different places' and *-uku* 'iterative in one direction'

The iterative markers *-bëkin* and *-uku* both indicate that the event is repeated several times together with specifying its distribution in space. They are different in that *-uku* is used when the repetitions of the event happen along a straight direction or following a predetermined path, while *-bëkin* is used when the repetitions are unsystematically distributed throughout space. One example of each suffix follows in (31) and (32).

(31) *Isa_Kuna ka-kë*   *a=n*   *ain*   *manë xo=n*   *cháxki-**bëkin**-kin*
Isa Kuna say-NOM 3SG=erg 3SG:gen metal stick=INS poke-ITER-S/A>A:SE
kaisa   *a*   *uni=n*   *baka kamabi xëxá*   *'imainun rara=n*
NAR:REP:3 that man=ERG river all   current and   ancestor=GEN
papa=kama anë-ru-akë-x-a
big=PL   name-up-REM.PST-3-NON.PROX
'A man called Isa Kunabu named all the rivers and creeks, even the big ones, poking his metal stick several times here and there.'

(32) *kwan-i*   *kaisa*   *xa*   *ki-**uku**-kian-akë-x-ín*
go-S/A>S(SE) NAR:REP:3 IDEO say:INTR**-ITER**-going:INTR-REM.PST-3-PROX
'It is said that, going, (he) went (down to the ground) making the noise *xa* several times.'

2.3.3 Reduplication

Reduplication of predicates is often used to express iterativity or plurality of an event. Reduplication of verbs in Kakataibo can apply over the whole verbal word, the stem (i.e., the root and its derivational suffixes) or just the root. More research is needed to determine whether or not there are systematic semantic differences associated with these different types of reduplication strategies, although differences in scope are often transparent among these various types of reduplication. Ex-

ample (33) illustrates a case of verbal reduplication applying to the whole word. The semantic interpretation of the reduplicated structure is clearly iterative.

(33) *ronru-tankëxun kaisa xëmën 'a-akë-x-a*
climb-S/A>A:PE NAR:REP:3 kinkajous:ABS kill-REM.PST-3-NON.PROX
*pia=n pia=n **'a-xun 'a-xun***
arrow=INS arrow=INS **kill-S/A>A kill-S/A>A**
ni-pat-akë-x-a
throw-down:TRAN-REM.PST-3-NON.PROX
'It is said that, after climbing, (he) killed kinkajous, killing and killing (them) with arrows, he threw down several.'

In the example of stem reduplication in (34) the form *ni-pat* 'throw-down' is reduplicated, but the switch-reference marker *-kin* 'S/A>A, simultaneous events' is not. The interpretation is again iterative.

(34) **ni-pat ni-pat**-*kin is-këx=bi kaisa*
throw-down:TRAN throw-down:TRAN-S/A>A:SE see-P>S=same NAR:REP:3
ain xanu 'a-kë uni a=x u-akë-x-a
3SG:GEN wife:ABS do-NMLZ man 3SG=S come-REM.PST-3-NON.PROX
'It is said that, throwing and throwing (animals from a tree), (the husband) saw the man who used to have sex with his wife coming.'

Finally, in (35), only the root *nëa* 'tie' is reduplicated, indicating that the process of tying was repeated but there was only one motion event associated with it: something like 'tie (something) several times while going upward only once'. If the reduplication had applied over both the root and the directionals (i.e., *nëarubian nëarubian)*), the translation would have been: 'tie (something) repeatedly while going upward several times (i.e. going upward tying up, coming downward, going upward tying up, and so on)'.

(35) *ukairi-o-xun kaisa **nëa nëa**-ru-bian-kin*
Ladder-FACT-S/A>A:SE NAR:REP:3 **tie tie**-up-going:TRA-S/A>A:SE
'a-bait-kin kaisa ka-akë-x-ín
do-DUR-S/A>A:SE NAR:REP:3 say-REM.PST-3-PROX
'It is said that, making a ladder, tying it several times while going upwards, doing it for a long time, he said ...'

3 Syntax and agreement

The encoding of plurality in Kakataibo is optional. Therefore, plurality may be encoded on the NP, on the verb or on both, and there is no obligatory NP-V number

agreement in the language. Indeed, plurality need not be expressed at all, as also seems to be the most common pattern for South American languages, in which plurality is optional (see Krasnoukhova 2012: 259 and Krasnoukhova, this volume). Number agreement is neither found within the NP.

4 Semantics and discourse

A detailed study of the discursive properties of number markers in Kakataibo is still to be conducted. What I present in this section is a list of facts that may contribute to a future exploration of this topic.

A first point to be made is that number is not an obligatory category in Kakataibo. Neither the nominal plural =*kama* nor the verbal plural -*kan* are mandatory in reference to plural entities. Plurality is often induced from the context without the presence of any plural marker. This opens the question of what triggers the use of a plural marker in Kakataibo discourse. Although disambiguation seems to be a good possible explanation, the topic requires a more detailed corpus study.

Other markers presented in this chapter trigger very specific pragmatic interpretations. For instance, Kakataibo speakers report that the marker -*ina(k)* 'generic' is "used to advise the young people". Such an explanation fits the semantic description proposed here, since "advising" takes the form of telling young people things like "you should be aware of snakes" or "our ancestors used to hunt tapirs and you should hunt tapirs as well" where "snakes" and "tapirs" are generic (Zariquiey 2018: 227).

The collective marker -*baë* is mainly used for kinship terms, but they can also be pluralized by means of the general plural =*kama*. For example, both *chichibaë* and *chichikama* 'grandmothers' are found in the corpus. While the former translates as something like 'female dead ancestors', the latter refers to the grandmother of the speaker and her sisters (who may be alive and can be listed).[4] Note that both forms can co-occur, like in *rara-baë=kama* 'ancestor-COL=PL'. If this happens, the plural enclitic always follows the collective suffix (Zariquiey 2018: 228). In turn, the events modified by -*rabat* 'distributive' are interpreted as being disorganized and chaotic, but this seems to be an implicature associated with certain types of events (like fighting).

5 Conclusions

In this chapter, I have put together the data available to me on how the category of number is encoded in Kakataibo. Most of the data presented come from previous

---

4 Note that in Kakataibo one's grandmothers and their sisters are equally considered "grandmothers".

descriptive studies of the language (particularly Zariquiey 2011 and 2018), but this is the first time that these data are presented in a systematic way to provide an overall account of the topic. I have shown that there are multiple manifestations of the category of number in Kakataibo: number in the nominal domain, number in the verbal domain and number in other lexical categories (numerals and quantifiers).

In the nominal domain there are four bound morphemes that express number in Kakataibo: the plural marker =*kama*, the distributive marker =*ti(i)bi*, the generic marker -*ina(k)* and the collective marker -*baë*. The generic and collective suffixes have very limited distributions, being associated with specific semantic classes of nouns. The distributive and, particularly, the plural marker have a much wider distribution and appear on almost all nouns, as well as on some types of NP post-head modifiers. A few nouns can be considered non-count nouns, and therefore cannot be pluralized (for instance, *uñe* 'rain' or *ni* 'jungle'). In addition, nouns carrying the generic suffix -*ina(k)* cannot be pluralized by means of =*kama* (but they can carry the collective derivational marker -*baë*).

There are significant intergenerational differences regarding number in personal pronouns. Elderly speakers establish a three-way number distinction on pronouns: singular, dual/paucal and plural, as well as a clusivity distinction in first-person plural pronominal forms. These number distinctions, however, are not attested among younger speakers. Since the general plural =*kama* is an enclitic (i.e. a phrasal modifier), there is no chance for NP-internal number agreement. The marker =*kama* will appear only on the rightmost element of a NP, which may be a noun or an NP modifier (since determiners, numerals, quantifiers and adjectives may appear as post-head NP modifiers).

Marking of plurality is not restricted to the NP. There is also a plural marker that attaches to the verb, -*kan*, which is used exclusively for third person subjects. There are also various other number-related verbal markers that have been discussed in this chapter: -*rabat* 'distributive', -*pat* 'plural objects', -*tabat* 'once', -*tëkën* 'again', and two iterative markers: the suffixes -*bëkin* and -*uku*, which express different types of distribution of the events in the space. Iteration (or event number) may also be expressed through verbal reduplication of various types (reduplication of the root, reduplication of the stem or reduplication of the whole word).

The encoding of plurality in the verbal word or NP is optional. Therefore, plurality may be encoded on the NP, on the verb or on both, and there is no obligatory NP-V number agreement in the language. Indeed, plurality need not be expressed at all, as also seems to be the most common pattern for South American languages.

What is fascinating about Kakataibo is the presence of a large list of number-related markers (both on the verb and on the noun), which are used for particular number configurations of entities and events. If we put all these different markers together, the category of number in Kakataibo certainly becomes a rich functional domain relevant to the general typology of number.

Acknowledgements

Abbreviations

>	interclausal switch-reference tracking (dependent > main). For example, 'P>S' indicates that the P argument of the dependent clause is the S argument of the matrix clause.
1	first person
2	second person
3	third person
A	most agentive argument of a transitive construction
ABS	absolutive
BEN	benefactive
CAUS	causative
COL	collective
COM	comitative
DES	desiderative
DIM	diminutive
DIST	distributive
DS/A	different subjects
DS/A/P	different subjects and objects
DUB	dubitative
DUR	durative
ERG	ergative
FACT	factitive
GEN	genitive
GENE	generic
HAB	habitual
HAR	harmonic
IDEO	ideophone
IMP	imperative
IPFV	imperfective
INS	instrument
INT	interrogative
INTR	intransitive
ITER	iterative
LOC	locative
NAR	narrative register
NEG	negative
NMLZ	nominalizer
NON.PROX	non-proximal to the addressee
NP	noun phrase
NUM	numeral
P	most patientive argument of a transitive construction
PA	participant agreement

PST	past
PE	previous dependent event
PFV	perfective
PL	plural
PROX	proximal to the addressee
REM.PST	remote past
REP	reportative
S	single argument intransitive construction
SE	simultaneous dependent event
SG	singular
TRAN	transitive

References

d'Ans, Marcel. 1973. Reclasificación de las lenguas pano y datos glotocronológicos para la etnohistoria de la Amazonía peruana. *Revista del Museo Nacional* XXXIX. 349–369.

Dixon, R.M.W & Alexandra Aikhenvald. 1999. Introduction. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian Languages*, 227–50. Cambridge: Cambridge University Press.

Krasnoukhova, Olga. 2012. *The noun phrase in the languages of South America*. Utrecht: LOT publications.

Loos, Eugene. 1999. Pano. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian Languages*, 227–50. Cambridge: Cambridge University Press.

Fleck, David W. 2013. Panoan languages and Linguistics. *Anthropological Papers of the American Museum of Natural History*, 99. New York: American Museum of Natural History.

Shell, Olive A. 1965. Pano Reconstruction. PhD dissertation. University of Pennsylvania, Philadelphia.

Tessmann, Günter. 1930. *Die Indianer Nordost-Perus: Grundlegende Forschungen für eine systematische Kulturkunde*. Hamburg: Friederichsen, de Gruyter & Co. m.b.H.

Valenzuela, Pilar M. 2010. Ergativity in Shipibo-Konibo. In Spike Gildea and Francesc Queixalós (eds.), *Ergativity in Amazonia*, 65–96. Amsterdam: John Benjamins.

Valenzuela, Pilar. Armonía transitiva en Pano y en Takana. In Antoine Guillaume and Pilar Valenzuela (eds.), Estudios sincrónicos y diacrónicos sobre lenguas Pano y Takana. *Amerindia* 39(2). 409–454.

Valle, Daniel & Roberto Zariquiey. 2019. Grammatical nominalization in Kakataibo. In Roberto Zariquiey, Masayoshi Shibatani & David Fleck (eds.), *Nominalization in languages of the Americas*, 517–538. Amsterdam: John Benjamins.

Zariquiey, Roberto. 2011a. Hacia una dialectología del idioma Kakataibo. *Lexis* 35(1). 5–46.

Zariquiey, Roberto. 2011b. Relaciones gramaticales en Kakataibo. In Willem Adelaar, Pilar Valenzuela Bismarck & Roberto Zariquiey Biondi (eds.), *Estudios sobre lenguas andinas y amazónicas. Homenaje a Rodolfo Cerrón-Palomino*, 327–247. Lima: PUCP.

Zariquiey, Roberto. 2013. Tessmann's <Nokamán>: a linguistic characterization of a mysterious Pano group. *Cadernos de Etnolinguistica* 5(2). 1–46.

Zariquiey, Roberto. 2016. Target, embedding and switch-reference in Kakataibo (Panoan, Perú). In Rik van Gijn, Rik & Jeremy Hammond (eds.), *Switch-reference 2.0*, 473–492. Amsterdam: John Benjamins.

Zariquiey, Roberto. 2018. *A grammar of Kakataibo*. (Mouton Grammar Library 75). Berlin: De Gruyter Mouton.

Zariquiey, Roberto & David W. Fleck. 2012. Body-Part Prefixation in Kashibo-Kakataibo: Synchronic or Diachronic Derivation? *IJAL* 78(3). 385–409.

Daniel Harbour and Andrew McKenzie
19 Grammatical simplexity: Number in Kiowa

Abstract: The Kiowa number system is deceptively simple. It has three numbers (singular, dual, and plural) and each noun morphologically makes at most a two-way distinction. Strikingly, though, nouns do not encode singular, dual, or plural. Instead, their basic forms cover a subset of these numbers, and a single morpheme, the 'inverse', is used for the rest. Inverse marking covers singular, nondual, or plural, depending on semantic properties of the noun. Further semantic properties of the noun are reflected in number agreement on the verb. Agreement distinguishes singular, dual, and plural. However, depending on the noun, singular or dual can be used in the plural, or vice versa, reflecting qualities such as animacy, collectivity, or internal complexity. Inverse-marked nouns generally require 'inverse agreement', but this too can be overridden on semantic grounds, especially for animates, some of which take special 'empathic agreement'. The most reliable marker of cardinality turns out to be verbal number, but it is encoded for only a small number of suppletive roots. In addition to describing this system in detail, we distinguish verbal number from distributivity and lay out some unusual morphological, semantic, and discourse-level properties of number in Kiowa.

1 Overview

The fascination of Kiowa[1] for the theory of linguistic number lies in the fact that it embeds a common number system (singular, dual, plural) in an uncommon design of noun class system (one intimately tied to the semantics of number) and chooses the verb, not the noun or its modifiers, as its main morphological locus. The complexity is compounded by its simplicity, in that the class system is configured with a bare minimum of class-specific morphemes. So, the story of Kiowa number is one of multipurpose morphemes in a complex interplay with noun class distributed over nouns and verbs.

---

1 Kiowa is a Kiowa-Tanoan language spoken in Oklahoma. It has no official orthography, and here we employ a working orthography, derived from Redbird et al. (1962), that employs standard IPA values for consonants but with *y* for palatal glides and inline *h* for aspiration. For instance, velar stops are *g*, *k*, *k'*, *kh*. Vowels are marked for high tone (*á*) and falling tone (*â*), with low tone unmarked (*a*). Phonemic nasality is marked with an underscore (*a̲*). Long vowels are written as doubled or as glides (*aa, ey, ii, ow, uu*) and diacritics are added only to the first element of such digraphs (long high nasal *a* is written *á̲a*). IPA vowel qualities are used, except that the low back rounded vowel ("open o") is written as short *au* and long *aw*.

Common nouns fall into eight classes. Nominal marking distinguishes four of these classes and does so with remarkable economy. One suffix, the so-called "inverse," marks singular ($_1$), plural ($_{3+}$), or both, depending on the noun (Wonderly, Gibson, and Kirk 1954). For instance, the inverse marker *-dáu* marks the plural of 'insect' but the singular of 'tree':[2]

(1) *pól*   'insect(s$_2$)'
     pów-dáu 'insect$_{3+}$'

     áa-dau 'tree'
     áa   'trees$_{2+}$'

The inverse never marks the dual ($_2$) in Kiowa. Non-inverse-marked forms are termed "basic."

The bulk of the number/class system is borne by verb inflection. Agreement prefixes distinguish singular, dual, and plural, which nouns themselves do not mark. However, this simple picture is blurred by complicating factors. First, inverse-marked nouns trigger "inverse agreement" that overrides semantic agreement irrespective of number of referents. Second, some noun classes trigger singular agreement for plural denotees, and others singular no matter the number. Third, morphological restrictions based on argument structure affect which numbers cooccur in the verbal agreement. Semantic number and morphological number are, therefore, sometimes obscured or mismatched.

Noun classification is semantic. Classes include motile objects (especially animates), collective inanimates, pluralia tantum, and nongranular mass nouns. The class-defining semantic criteria connect nonarbitrarily to the classes' morphological signatures. For instance, collectives trigger singular agreement in the plural, and pluralia tantum, plural agreement even when nonplural; and in some of its uses, inverse marking is reminiscent of a singulative. Doublets, though rare, give extra insight into the organisational principles.

Even though noun class and number are so exuberant in Kiowa, Harbour (2008) shows that number marking is associated with determiners and similar functional projections in the syntax. Marking occurs with demonstratives, 'only', and in possessives, as well as in verbs and with relative clauses. Bare nominals can bear number marking (2), but only as part of a full DP: noun roots are number neutral when compounded or incorporated (3).

---

2 The example shows two regular morphophonological changes: the suffix *-dáu* triggers deletion of root-final *l* with compensatory lengthening of the root vowel (*pól* becomes *pów*), and acquires low tone after the lexical item *áa*.

(2) *táa-de*
eye-BAS
'eye(s$_2$)'

táa-gau
eye-INV
'eyes$_{3+}$'

(3) *táa(*de/*gau)-khạwgya*
eye(<u>BAS</u>/<u>INV</u>)-skin
'eyelid(s$_{2+}$)'

Né-   táa(*de)- k'ạwde.
:1SD:3DU-eye(<u>BAS</u>)-be bad
'I have poor eyesight.'

With the exception of verb agreement and verbal number, these categories distinguish only inverse and "basic" (noninverse) and they do so with a high degree of syncretism: demonstratives, 'only', and relative clauses all use the same marking, as do some possessives. Nouns use the same marking but only if none of the rich array of phonologically sensitive allomorphs is called for. 'Only' is further interesting as a rare locus of optional number marking. Kiowa number marking is otherwise obligatory.

Verbal number exists but is not prominent in Kiowa. Suppletion of certain verbs reflects participant number (as defined in Corbett 2000), while their adverbials reflect event number. Spatial distributives indirectly reflect event number by entailing significant plurality.

2 Pronominal, nominal, and verbal number

2.1 Generalities

At first glance, pronominal number in Kiowa stands apart from nominal and verbal number in that nouns and verbs have intricate number morphology whereas number on personal pronouns is wholly absent. However, this is a misreading of the system. Pronominal and nominal number are in fact alike and verbal number as a distinct semantic category is peripheral. This situation arises because number in Kiowa is inextricably linked with class and both these nominal properties are primarily expressed via agreement on the verb. The verb, by contrast, does not show number independently of its arguments (verbal number is highly limited, as per the previous paragraph). Verbs reveal that personal pronouns, though uninflected for number, are almost as well-endowed for number as common nouns.

The following subsections address pronominal (sections 2.2), nominal (section 2.3), and verbal (section 2.4) number. The system that emerges is striking in several respects. First, Kiowa nouns do not mark number in the strict sense. Rather, nouns show a two-way contrast, basic versus inverse, which amalgamates number with noun class. Second, number and class are dislocated. Though they belong semantically to the noun, they are expressed primarily on the verb. Agreement reveals three semantic numbers (singular, dual, and plural) and eight noun classes, much of which is undetectable on nouns. Third, the noun classes are only fully encoded via the number system, recycling number categories to differentiate between animates and inanimates, group and nongroup plurals, and pluralia tantum, granular and nongranular mass nouns. Fourth, the resulting system runs against the animacy hierarchy (Corbett 2000) in being simpler for first person than for third, with second person shifting allegiance depending on how one counts distinctions.

We refer (atheoretically) to argument indexing on the verb as agreement. We notate intransitive agreement as *z-*, transitive as *x:z-*, ditransitive as *x:y:z-*, and intransitive with an experiencer as *:y:z-*, where *x* is the agent, *y*, the experiencer or oblique, and *z*, the internal argument (transitive object or unaccusative subject).

Relative to the themes of this volume, we note that Kiowa lacks number words (e.g., no word meaning PL), classifiers, nominal case (though agreement registers argument roles), and marking of definiteness (other than in demonstratives and the like).

2.2 Pronominal number

There are only two personal pronouns in Kiowa: *nâw* for all first persons, irrespective of number and clusivity, and *ám* for all second persons. This simplicity contrasts with the complexity of verbal agreement, which encodes three argument types (agent, indirect object, direct object) while distinguishing four morphological numbers (singular, dual, plural, inverse) and up to four persons (inclusive, exclusive, second, third; as well as a special animate plural (81)). We lay out these contrasts and show they also apply to the reflexive pronoun *ą̂wgau* 'self'. We then compare first and second person pronouns with demonstratives, which Kiowa employs in place of third person pronouns, and which bear number based on the nouns they replace. We end with a brief argument that the pronouns can be conjoined and do not enter in comitative constructions.

Second person (4) illustrates the difference in number sensitivity between pronouns and agreement. Pronominal *ám* is invariant, while agreement encodes singular (*em-*), dual (*ma-*), and plural (*ba-*):

(4) *Ám em- dáw.*
    2   2SG-be
    'You$_1$ are.'

Ám ma- dáw.
2   2DU-be
'You$_2$ are.'

Ám ba- dáw.
2   2PL-be
'You$_{3+}$ are.'

First person *náw* is also invariant but its agreement makes only a twofold number distinction, singular and nonsingular (see table 3 for relationship to the animacy hierarchy). This is shown for exclusive (*a-* versus *e-*) and inclusive (*ba-*) in (5) and for general first person (*ę́-* versus *dáu-*) in (6):[3]

(5)   *Náw a- dáw.*
      1   1SG-be
      'I am.'

      Náw e- dáw.
      1   1EX.NSG-be
      We.EX$_{2+}$ are.'

      Náw ba- dáw.
      1   1IN.NSG-be
      'We.IN$_{2+}$ are.'

(6)   *Náw ę́- góp.*
      1   2/3.SG:1SG-hit.PFV
      'You$_1$/he/she hit me.'

      Náw dáu- góp.
      1   2/3:1NSG-hit.PFV
      'You/he/she/they hit us$_{2+}$.'

First person dual and plural are distinguished only by a small class of suppletive verbs (table 4). For these verbs, a single root form (here, *ą́agya*) is shared by singular (7) and dual (top examples of (8)–(9)), another (here, *k'úl*) by the plural (bottom sentences), irrespective of clusivity.

(7)   *Náw a- ą́agya.*
      1   1SG-seated.SG/DU
      'I am sitting.'

---

3 Clusives are restricted to agents and intransitive subjects. Indirect and direct objects have only general first person.

(8) *Náw e-   áagya.*
    1   1EX.NSG-seated.SG/DU
    'We.EX$_2$ are sitting.'

     Náw e-   k'úl.
    1   1EX.NSG-seated.PL
    'We.EX$_{3+}$ are sitting.'

(9) *Náw ba-   áagya.*
    1   1IN.NSG-seated.SG/DU
    'We.IN$_2$ are sitting.'

     Náw ba-   k'úl.
    1   1IN.NSG-seated.PL
    'We.IN$_{3+}$ are sitting.'

Derivatives of first and second pronouns formed with the bound stem *-hįi* 'genuine, original' also lack inverse marking: *náwhįi* and *ámhįi* invariantly cover all numbers. This holds even though *-hįi* itself does take inverse marking with other nouns, as in *áa-hyǫy* 'cottonwood', *áa-hįi* 'cottonwoods$_{2+}$'.

    Likewise, the all-person reflexive/emphatic *ą́wgau* refers to any number. It also refers to any person.

(10) *Ą́wgau a- khą́waun né gyat- hémbását.*
     self   1SG-poor   but 1SG:3PL-provide for.PFV
     'I am poor but have provided for myself.'
     (Harrington 1946: 238)

(11) *Ą́wgau ba- bówow-gau   ówpêy ba- thǫw-t'éykyá- dey- yau.*
     self   2PL-inept-   and.SAME LOC 2PL-leg- strained-stand-DISTR
     'You are inept, standing along there with strained legs.'
     (P. McKenzie n. d.-b)

(12) *Ą́wgau tsâw dáu-   tét.*
     self   thus 3SG:1NSG-tell.PFV
     'He himself told us.'
     (Toyebo 1957c)

Kiowa lacks third person pronouns, deploying demonstratives instead. These distinguish number via inverse marking, unlike *náw*, *ám*, and *ą́wgau*. The inverse is present when it would be present on the head noun, whether or not the noun is overt. In (13a), basic *éyde* reflects the basic head noun *k'âw*, while in (13b) the demonstrative bears the inverse.

(13) a. *Ę́y- de (k'âw) Ø- gúlkháun-daw.*
PROX-BAS knife 3SG-burn.NV- be
'This (knife) is burnt.'

  b. *Ę́y- gau (k'âw-gau) e- gúlkháun-daw.*
PROX-INV knife-INV 3INV-burn.NV- be
'These (knives) are burnt.'

Third person categories are discussed in sections 2.3 and 4.

 Given the absence of number on personal pronouns, form alone does not determine whether *náw* in a conjunction such as *Mary gau náw* is singular ('Mary and I') or plural ('Mary and us'), a feature of the questionnaire underlying this volume. Weighing against the latter, comitative analysis, the connective *gau* is identical to that used for both (14) nominal and (15) some clausal conjunction (as part of the switch reference system; Watkins 1984: 236–241; A. McKenzie 2011, 2012):

(14) *Pêysa̱adau gau phola̱ayop hegáu hę́y gya-dáw.*
quail.INV and rabbit.INV then PRIV 3PL-be
'Quails and rabbits are no more.'

(15) *Tségun Ø- a̱adêy gau em- bét- tawyii.*
dog 3SG-sit.IPFV.EVID and.SAME 3SG:RX-bark-act.IPFV.EVID
'The dog was sitting and barking.'
(Toyebo 1957b)

Nonetheless, it is common for conjunctions to split with only one element of the conjunction agreeing. In (16), only Big Tree agrees with the verb, or in (17), the only object agreement is with *hólda* (which is always plural).

(16) *Pha̱aow Káuy- gú á- dáw-gau gyá- pén- mau Bétsép Ø-*
three Kiowa-INV 3AN.PL-be- INV 3AN.PL:3PL-butcher-IPFV Big Tree 3SG-
dáw-gau K'ówdebo̱hon gau Sę́ygauy.
be- and Gotebo and Saingko
'Three Kiowas present were butchering: there was Big Tree, and Gotebo and Saingko.'
(McKenzie n. d.-f)

(17) *Hólda ya̱- a̱w- aw gau k'aunbo̱ho̱w-dau gau*
Shirt 2SG:1SG:3PL-temporarily-give.IMP and hat- INV and
tháydeholda gau tokį̄inį́i.
coat and boots
'Lend me your shirt and hat and coat and boots.'
(Spotted Horse 1957)

Conjunctions are further discussed in (93)–(95).

Some *wh*-pronouns are sensitive to number, reflecting the number of the answer the speaker anticipates. For (19), a plural answer is expected, whereas (18) need not carry any such expectation.

(18) *Hâundé a-   bǫ́w?*
    who.<u>BAS</u> 2SG:3SG-see.PFV
    'Who did you see?'

(19) *Hâungáu be-   bǫ́w?*
    who.<u>INV</u> 2SG:3AN.PL-see.PFV
    'Who did you see?'
    (Laurel J. Watkins, p.c.)

The inverse-marked plural *wh*-word does not require an exhaustive answer.

We now turn to the number system for common nouns, which underlies the demonstratives discussed above.

2.3 Nominal number

This section introduces inverse marking and its relationship to noun classes. A selection of doublets illustrates the semantic basis of the four classes of inverse marking. Allomorphs of inverse marking admit little free variation and do not vary by noun class. Indeed, verb agreement is a truer indicator of noun class: not only is it more accurate for the four inverse-based classes, but it reveals four more, centred on noncanonical uses of singular and plural agreement. In this context, we discuss group plurals, pluralia tantum, mass nouns, and abstract nouns, which include nominalisations and complement clauses. We end by showing that plural is the default number in Kiowa and that, in light of nominal patterns, pronominal number (section 2.2) runs against crosslinguistic tendencies surrounding animacy and number.

Kiowa nouns are not marked directly for semantic number. Instead, they have, at most, two forms, "basic" and "inverse". Basic nouns are usually unmarked, and inverse nouns are almost always marked. The number these forms denote varies according to the noun's class, as shown in table 1. Basic nouns trigger semantically transparent agreement (barring semantically grounded exceptions discussed later), while inverse nouns trigger inverse agreement. We refer to the classes via acronyms reflecting the agreement pattern they trigger on the verb (following Harbour 2008). An SDL noun triggers singular agreement (S) in the singular, dual agreement (D) in the dual, and inverse (I) in the plural. An IDP noun triggers inverse in the singular, dual in the dual, and plural in the plural, and so forth. For all classes, any noninverse form is basic.

Tab. 1: Four classes of inverse marking.

	SDI 'horse(s)'	IDI 'orange(s)'	IDP 'baked good(s)'	SDP 'mattress(es)'
singular	*tsêy*	*thówt'ólạw-gau*	*éy-gáu*	*pálts'e*
dual	*tsêy*	*thówt'ólạw*	*éy*	*pálts'e*
plural	*tsêy-gau*	*thówt'ólạw-gau*	*éy*	*pálts'e*

Of the classes in table 1, SDI and IDP are the most populous. The central property of SDI nouns is motility, and the class subsumes nearly all animates, other mobile objects (sun, moon, cars, wheels), mobile body parts and major internal organs (leg, knee, heart, liver), and various implements, particularly for cutting (awl, knife, spoon). Many inanimates, by contrast, are IDP nouns, including many plants (sage, sunflowers, weeds) and their parts (sticks, leaves, beans), as well as relatively immobile body parts (bones, noses, udders) and implements (bed sheets, smoking pipes, arrowheads). SDP nouns are apparently less numerous than IDP. The predominant members of the class appear to be clothing (shoes, boots, rings), but some natural objects (bark, stones) are included, as is meat when portioned. IDI nouns are rare, being restricted to midsized fruit (apples, oranges, plums, tomatoes) and hair and eyebrows (the rest of the face and body have 'fur').

Doublets, though neither abundant nor highly systematic, shed light on the organising principles. For instance, animal tails and similar parts, involving the stem *thón* 'tail', are classified according to whether they are internally motile (SDI *tsêython* 'horsetail', *pówthón* 'beaver tail') or rigid (IDP *kúutothon* 'bird tail', *phíithón* 'porcupine quills'). In a similar vein, *t'áulthaun* is IDP when it means 'beans' but SDI, like other major organs, when it means 'kidney'. The term *kháuy* can be applied to cloth, rags and hide. Generally, compounds built on this stem are SDP (*tábekhauy* 'blindfold', *máunkháuy* 'gloves', *ts'ówk'íkhauy* 'slingshot'). However, when applied to hides, such terms are SDI, like animals (*t'ápkháuy* 'deer hide', *t'aupkháuy* 'buckskin').

Number suffixes vary substantially in form (table 2). Most basic nouns have no marker. Some have the suffix *-de*, which also appears with demonstratives and relativizers. A handful have *-da* or *-gya*. These are all lexically idiosyncratic. For inverse markers, lexical idiosyncrasy is limited and the main determinant is phonology. Class is not a determining factor. Pairs in table 2 demonstrate this clearly. A full list of inverse forms can be found in Watkins (1984:80) or Harbour (2008:55).

Only rarely is there free variation between suffixes. An interesting case is *tháa* 'wife', for which the inverse can be suffixal *tháa-gau* 'wives' (wife-INV) or root-modifying *thêy* (wife.INV). A productive suffix showing variation is *-hịi* 'genuine, original' (section 2.2), which regularly alternates between two inverse forms, as in SDI *tsêyhyọy/tsêy-hyop* 'dogs$_{3+}$ from *tsêyhịi* 'dog(s$_2$)'. There is also dialectal variation. The great Kiowa linguist Parker McKenzie could tell where a speaker was from by their inverse form of *éy* 'bread': *éybáu*, *éybáut*, *éygáu*, or *éygáut* (Watkins and Harbour 2010).

Tab. 2: Selection of basic and inverse suffixes.

Noun	Class	*Basic*	*Inverse*
'insect'	SDI	*pól*	*pów-dáu*
'hair'	IDI	*ául*	*áw-dáu*
'friend'	SDI	*kóm*	*kǫ́w-bau*
'student'	SDI	*mauthêm*	*mauthêy-bau*
'eye'	SDI	*táa-de*	*táa-gau*
'loincloth'	IDP	*tâuy-de*	*tâuy-gop*
'flower'	IDS	*áakhįį-gya*	*áakhįį-gaut*
'child'	SDI	*iip'áw-gya*	*iip'áw-gaut*
'arrow'	IDS	*zéyba*	*zéybaut*
'chief'	SDI	*k'yátáy*	*k'yátây*
'deer'	SDI	*t'áp*	*t'áp*
'mattress'	SDP	*pálts'e*	–
'life'	SSS	*k'yákôm-da*	–

For all of its complexity, inverse marking provides an insufficient basis for defining noun classes. A simple reason is that some nouns trigger inverse agreement without taking an inverse suffix. For instance, compare (20) SDP *pálts'e* 'mattress' and (22) SDI *k'âw* 'knife' with (21) SDI *t'áp* 'deer'.

(20) *pálts'e*   Ø-   *dáw*
      mattress 3SG-be
      'It's a mattress'

  pálts'e   *ę-*   *dáw*
  mattress 3DU-be
  'It's two mattresses'

  pálts'e   *gya-dáw*
  mattress 3PL-be
  'It's some mattresses'

(21) *t'áp* Ø- *dáw*
      deer 3SG-be
      'It's a deer'

  t'áp *ę-* *dáw*
  deer 3DU-be
  'It's two deer'

  t'áp *e-* *dáw*
  deer 3INV-be
  'It's some deer'[4]

---

4 The form with the plural (*t'áp gya-dáw*) is grammatical, but gives an existential reading 'there are deer; deer are available'.

(22) *k'âw   Ø-   dáw*
    knife 3SG-be
    'It's a knife'

    k'âw   e̜-   dáw
    knife 3DU-be
    'It's two knives'

    k'âw-gau e-   dáw
    knife-INV 3INV-be
    'It's some knives'

In terms of absence of suffixation, *t'ap* aligns with *pálts'e*. However, in terms of agreement, it aligns with *k'âw*: for the plural, both trigger inverse agreement (*e-*), rather than plural (*gya-*). Four factors urge us to consider *t'ap* as an SDI noun with a null inverse. First, regarding 'deer' as SDI brings it into the same class as other animates. Second, zero inverse is phonologically well defined: it occurs mostly after *p/t*. Third, overt inverse reemerges if *t'áp* is not word final. The compound *t'áp-kho̜wgya* [deer-black] 'black deer$_{1/2}$' ends in *-gya* for singular and dual, but switches to *t'áp-kho̜wgaut* 'black deer$_{3+}$', with an overt inverse, for plural referents (cf, *iip'áw-gya* 'baby, babies$_2$', *iip'áwgaut* 'babies$_{3+}$', table 2). Fourth, and most substantively, agreement is a much truer reflection of noun class than noun marking is.

The inverse suffix defines only four classes (table 1). However, common nouns fall into eight. For inanimates, nouns that generally form collections (of various sorts) take singular rather plural agreement for plural reference. This splits nouns that take inverse in the singular into two classes, IDP and IDS, and nouns that take no inverse into SDP and SDS. The IDS class is reasonably well populated, including mountains and large trees, whereas SDS is rather small, comprising clouds, paths, some bodies of water, and other things without predictable shapes as well as things that work for joint effect (fingers, lamps; cf, IDS 'guns' in (24)). Doublets again give a sense of the difference: *áa* means 'stick' if IDP but 'tree' if IDS; *tów* means 'tipi' if PPP but 'house' if SDS. (See Watkins 1984:86–87) for differences in compounds built on *áa* 'tree, stick', and (72)–(73) for 'finger, arm'.

Though Kiowa has no singulars with intrinsically plural reference, like English 'police', singular agreement is an occasional option for animates, giving a group-like reading, as is explicit in the translation of (24).

(23) *T'áukháuy-gú bá-   paw-kaun- taw.*
    mules-   INV 2PL:3SG-lead-bring-MOD
    'You will bring the mules.'
    (P. McKenzie n. d.-a)

(24) *Sóley- gau Ø- p'éyde-hel gigáu há̠wgya Ø- dów-*
soldier-INV 3SG-appear-EVID and then.SAME guns 3SG:3SG-hold-
dêy gigáu Ø- kaum-hâape-hel.
IPFV.EVID and then.SAME 3SG:3SG-aim- raise- EVID
'A group of soldiers appeared and they had guns and aimed them.'
(P. McKenzie n. d.-a)

Despite their singular agreement, SDI 'mules' and 'soldiers' retain the inverse marking necessary to their semantic plurality (in contrast to IDI nouns, (39)).

Words for collections, like 'flock' or 'shoal', are (to our knowledge) unattested, except for *káutâw* 'livestock, (horse/cattle) herd'. Etymologically, the word is related to SDI *kául* '(beef) cow, buffalo'. Irregularly for a noun in -*l*, *kául* takes a zero inverse. *Káutâw* is almost the expected inverse (cf, *ául* 'hairs₂', *áwdáu* 'hair(s₃₊)'; *tâl* 'skunk(s₂)', *táttau* 'skunks₃₊'). So, it is likely an archaic inverse (with short root syllable, voiceless *t*, and long falling-tone suffix) that has survived by dint of lexically unique semantic drift to a collective reading.

Pluralia tantum, abstract, and mass nouns also differ in ways that are only apparent from agreement. Pluralia tantum nouns like *hólda* 'shirt' and *aat'auhâui* 'war bonnet' and abstract nouns like *khá̠wgya* 'name' and *tó̠wgyá* 'word, speech, language' trigger plural agreement, irrespective of the number referred. Kiowa pluralia tantum nouns can be counted directly as units, as other languages also allow (Comrie 2001; Doron and Müller 2014; Lima 2014).

(25) *Páagau / yíi / háote hólda gyat- háwgya.*
one two several shirt 1SG:3PL-get.PFV
'I bought one/two/several shirts.'

Nongranular mass nouns, like *thó̠waulkhauy* 'whisky' and *t'eléppenha̠a* 'honey', display the opposite behaviour, triggering singular agreement even if made into portions (26). One might regard these nouns as *singularia tantum*.

(26) *Yíi thó̠waulkhauy gya- thóm / *nen- thóm.*
Two whisky 1SG:3SG-drink.PFV 1SG:3DU-drink.PFV
'I drank two whiskies.'
(Harbour 2008:29)

(Granular mass nouns agree like pluralia tantum nouns, but they permit inverse marking for grain readings. For instance, *péygya* 'sand' is ambiguous between a PPP mass noun with no inverse, and an IDP count noun 'grain of sand' with inverse *péygaut*. Harrington (1928, *passim*) offers many examples.)

Nominalisation is not a prominent process in Kiowa grammar. Deverbal nominals such as *îlgya* 'act of admonishing' (from *îl* 'admonish') and *k'íithêmgya* 'act of

gathering wood' (from *k'íithêm* 'gather wood') exist, but one does not encounter, for instance, 'the teacher's admonishing of the students'.[5] Ending in *-gya*, these nouns are abstract PPP nouns.

(27) *K'ombââlgya háun an gya- áum-g- âw.*
    Imitating NEG HAB :3AN.PL:3PL-do- DETR-NEG
    'Imitation is hard to pull off.' (Harrington 1946:241)

A few predicates allow a personified nominal in *-k'íi* 'male'. Alongside *k'ómgyá* 'old age', for example, there is *k'ómk'íi* with the same meaning. Kiowa generally resists inanimate agents, but personified nominals escape this restriction:

(28) *K'óm-k'íi hegáu yá̱- dâate- dow.*
    old- MALE then 3SG:1SG:3PL-overtake-AUX
    'Old age is about to overtake me.'
    (P. McKenzie 1990)

Complement clauses also look like nominalisations. They bear basic number marking (*-de*) and, like the abstract nouns above, govern plural agreement (*án-* in (29), encoding also the third singular knowledge holder as the applicative):

(29) *Háun-an êlk'yoy gyát- sém- há̱wnâw- de án- khą̣uâwn-*
    NEG- HAB old man.INV :1NSG:3PL-secret-exhaust.NEG-BAS :3SG:3PL-pitiable-
    haygya-daw-do.
    known- be- because
    'Because she knows that sadly us old men don't give up our desires.'
    (P. McKenzie 1963a)

These closely resemble relative clauses, with the difference that relatives take basic/inverse marking from their head nouns. (30) is dual and hence basic marked (*-de*), (31) animate plural and hence inverse marked (*-gau*).

---

5 Story titles and picture descriptions use subordinate clauses. Low tone on 'story' and 'picture' in (i)–(ii) shows that these form a prosodic unit with the preceding verb and basic number marker.

(i) *Áadaual-kya Hobêy Ø- kówba- de- heytegya.*
    barrel- in Jack Wolf 3SG-disappear-BAS-story
    'The story of Jack Wolf disappearing into a barrel.'
    (P. McKenzie n. d.-b)

(ii) *áugau Dawk'yaíi Ø- iip'awgya-tsán- de- kut.*
    REL Christ 3SG-baby- arrive-BAS-picture
    'Pictures of Christ's birth.'
    (P. McKenzie n. d.-h)

(30) [*Pháw- gu̱u e̱-   dôw*]-*de e-   tséy-   hel*.
    buffalo-horn 3SG:3DU-hold-BAS 3SG:3DU-put on.NPL-EVID
    'He put on the buffalo headpiece that he keeps around.'
    (P. McKenzie 1963b)

(31) [*Maayóp áugau ét-   pi̱i- a̱wmau*]- *gau máw e-   déygyay*.
    woman.INV REL   3INV:3PL-food-make.IPFV-INV like   3INV-pour.DETR.PFV
    'The women who were cooking poured out.'
    (P. McKenzie n. d.-d)

The plural agreement triggered by complement clauses and abstract nouns alike points to plural as the default number in Kiowa. Consistent with this, plural agreement is used for dummy objects of unergatives (the subjects of which are external arguments syntactically), objectless experiencers, weather predicates, and some existential predicates:

(32) *Gyat-   ây*.
    1SG:3PL-run off.PFV
    'I ran off.'

    Án-   t'áudep.
    :3SG:3PL-kind
    'He/she is kind.'

(33) *Gya-sál*.
    3PL- hot
    'It's hot.'

    A̱wgáwpi̱i-gau gya-káun-   hel.
    buffalo-   INV 3PL-numerous-EVID
    'There were a lot of buffalo.'

Predication of absence with privative *héy* also uses plural agreement, overriding the singular agreement otherwise expected for 'calf' in (34a) (and the inverse expected for 'quails and rabbits' in (14)):

(34) a. *Ts'álii hegáu Ø   dą́wmêy*.
       calf   then   3SG-be.EVID
       'The calf was there.'

    b. *Ts'álii hegáu hę́y   gya-dą́wmêy*.
       calf   then   PRIV 3PL-be.EVID
       'The calf was gone.'
       (P. McKenzie n. d.-e)

In (35), situational plural agreement occurs with a suppletive predicate which appears in its singular form:

(35) *Háun gya-kyǫ́y-   gâw-t'aw negáu   Káuy- tǫ- gya héy gya-dáw-*
    NEG 3PL-long.SG-NEG-MOD and then.DIFF Kiowa-say-BAS PRIV 3PL-be-
    t'áw.
    MOD
    'It won't be long before the Kiowa language is gone.'
    (P. McKenzie 1993)

Suppletion and agreement mismatches are discussed in section 2.4.

Plural agreement is also generally used for unspecified null objects in Kiowa (36). However, this is overridden if a verb prototypically takes an object class that does not take plural agreement. For instance, 'drink' takes singular object agreement because liquids are sss (37).

(36) *Gya- bǫ́w-   hêl nau   ówgau hegáu éythâl Ø- k'íiáa-dǫwmey*
    3S:3P-look at.PFV-EVID and.DIFF far off then   corn   3SG-grow- be.EVID
    'He looked around and for a very long distance, the corn had grown.'
    (Wolfe 1957)

(37) *Gya-   thónmau gyat- gúttau-   de- tso*
    1SG:3SG-drink.IPFV 1SG:3PL-write.IPFV-BAS-as
    'I was drinking as I was writing.'

The existence of an animate class related to number marking recalls the animacy hierarchy. Corbett (2000:90) observes that number distinctions monotonically decrease as one descends the hierarchy. Crosslinguistically, first person tends to have at least as many number distinctions as second, which in turn has at least as many as third person, and so on for subsets of third person down the hierarchy. Kiowa does not conform to the hierarchy either in terms of full nouns or pronouns, or agreement.

First and second person pronouns in Kiowa have one form for all numbers (*náw* and *ám*, respectively). However, third person nouns and demonstratives can have two distinctive forms, basic and inverse, in violation of the crosslinguistic tendency. The difference is visible in the 'noun/pronoun' row of Table 3.

Tab. 3: Distinctive number forms and the animacy hierarchy.

	first	>	second	>	third	>	kin	>	human	>	animate	>	inanimate	
Noun/pronoun			1	|					2				|	1
Agreement prefix	2	|							3				|	1

Agreement, the morphological locus of noun class, reduces this exceptionality, but does not eliminate it. First person makes only a two-way distinction in number agreement (5)–(8). Second person (4), by contrast, makes the same three-way distinction as other animates, singular/dual/plural (21). So, a two-way contrast tops the hierarchy ('2' in the table) over a three-way one in the central segment ('3'). (Only with suppletion is there a three-way contrast for first person; (7)–(8).)[6]

In sum, Kiowa achieves a remarkable amount of nominal classification (four classes) via a single noun class suffix, the inverse. However, verb agreement is the central hub of number and class, revealing eight classes and various nuanced readings (as for groups of animates). Though some of the classes are minor (IDI, SDS), others are very large (SDI, IDP) and are fed by grammatical processes like nominalisation (PPP).

2.4 Verbal number

Verbal number in the sense of number morphology, beyond agreement, on the verb is relatively limited as a semantic category in Kiowa. Corbett (2000) divides verbal number marking into event number and participant number. Event number counts the actual occurrences that have the predicate's property, while participant number reflects the count of nominal arguments. Kiowa verbs express both: a limited set of suppletive predicates express participant number, and spatial distributives indirectly indicate event plurality.

Table 4 shows that suppletive predicates in Kiowa come in two classes. Four adjective-like predicates sensitive to (non)singularity, and seven verbs (plus their

Tab. 4: Suppletive predicates in Kiowa.

Root	(Derivative)	SG	(SG/DU)	DU	(DU/PL)	PL	
big		êl			bîn		
small		syáun			syân		
long		kyǫ́y			kị̄iníi		
short		tséy			tsáadów		
be sitting		ǫ́agya			k'úl		
be lying		k'áw			k'úl		
wander		thów			zém		
set/put	(be sat/set.INAN)	tséy	(tsél)			sáw	(sául)
lay	(land, fall against)	ts'ów	(ts'óygyá)			k'úu	(k'úygyá)
drop	(be dropped, fall)	ól	(ótkyá)			p'él	(p'étkyá)
sever, trim	(be cut, come apart)	t'ál	(t'átkyá)			tháa	(thátkyá)

---

6 We leave aside animate plural agreement (section 4), as it is not a different numerosity and because it would count as a minor (hence exceptional) number in Corbett's terms.

intransitive derivatives), sensitive to non(plurality). The choice of predicate tracks the participant number of the internal argument, irrespective of complications in agreement related to noun class. Hence, *álawgau*, the inverse of IDI 'apple', occurs with a singular suppletive when referring to one apple (38a) but with a nonsingular when referring to a plurality (38b).

(38) a. *Álaw- gau dáut-   syáun.*
        apple-INV :1NSG:<u>3INV</u>-small.<u>SG</u>
        'Our apple is small.'

    b. *Álaw- gau dáut-   syân.*
        apple-INV :1NSG:<u>3INV</u>-small.<u>NSG</u>
        'Our apples₃₊ are small.'

The same noun permits singular agreement (like an IDS noun) when it refers to different kinds of apples. Notwithstanding the semantically plural noun triggering singular agreement, suppletion reflects the plurality of apples (39). These readings indicate three or more kinds of apple (Watkins 1984:88).

(39) *Álaw dáu-   syân.*
     apple :1NSG:<u>3SG</u>-small.<u>NSG</u>
     'Our [kinds of] apples are small.'

Another instance of singular agreement with plural 'trees' and a plural-sensitive predicate can be found later in (51). For more detail on the suppletive system, including some unexpected patterns, see Harbour (2008: ch. 4).

Adverbials derived from singular-sensitive verbs are sensitive to verbal number rather than participant number. In (40), the singular-based adverbs qualify the event of giving as a whole, whereas the nonsingular-based ones (41) qualify each giving event in a plurality.

(40) *Ét-   te / syáun- dé gya- áw- mau.*
     big.<u>SG</u>-ADV   small.<u>SG</u>-ADV 3SG:3PL-give-IPFV
     'She is giving away a lot / a little.'

(41) *Bîn-   de / syân-   de gya- áw- mau.*
     big.<u>NSG</u>-ADV   small.<u>NSG</u>-ADV 3SG:3PL-give-IPFV
     'She is giving away a lot/little at a time.'

In (42), the singular qualifies duration, the nonsingular, each step of the stitching.

(42) *Tséy-   dé / tsáadów- té an em- sôu- gu.*
     short.<u>SG</u>-ADV   short.<u>NSG</u>-ADV HAB 3SG:RX-sew-IPFV
     'She sews for a short time / in small stitches.'

However, whatever the basis of the difference is, the roots, when used as predicates, only supplete for participant number (which might reflect the individual-level nature of the singular-sensitive predicates: one cannot, for instance, be short on multiple occasions in a single situation).

The second potential source of verbal number is distributives, which indicate that an event occurs multiple times, spatially distributed around an area.[7]

(43) *Bímkháuy-gya hê̱ygya gyat-   sáw-   gô- m.*
    bag-   in toy   3SG:3PL-put in.PL.OBJ-DISTR-PFV
    'She went around putting toys in the bags.'
    (A. McKenzie 2020)

Unlike argument agreement, which is prefixal, distributives belong to the suffix chain of the verb, coming after markers of transitivity (44) and before aspect (44)–(45).

(44) *Háundé   yá̱-   áum-dé- go- m.*
    something :1SG:3PL-do- DETR-DISTR-PFV
    'I am able to get things done.'
    (Redbird et al. 1962: §52)

(45) *Hágyây-   zol- ku táa-gau bét-   hot- gû̱u- yi̱i.*
    which.INDEF-vomit-to eye-INV :3INV:3INV-travel-DISTR-IPFV.EVID
    'Their eyes roved over the vomit piles.'
    (Harrington 1946:241)

A. McKenzie (2020) demonstrates that distributives denote a sum of atomic sub-events of the predicate's event argument, strewn about that argument's location. While the sum condition in theory allows duals, the strewing about entails a significant plurality or mass and asserts some kind of spacing.

Distributives' plurality is independent of participant number, as they may occur with singular arguments, as (46) shows for an intransitive and (47) for both subject and object of a transitive verb:

---

7 Distributives are distinct from habituals, a point raised by the volume questionnaire. For statives, habituals often bear no marking, as in (85). For active predicates, the habitual particle *an* is used, as in (27), (29), and (81). Distributive habituals can deploy both devices at once.

(i) *Béthaw kâwgaw- al k'ó̱wbau tsâw an á-   hot- gû̱u- yi̱i.*
    MIR   other.INV-also elder.INV thus HAB 3AN.PL-travel-DISTR-IPFV.EVID
    'We didn't realize that other old people ran around like that.'
    (Harbour, Watkins, and Adger 2012: 124)

Distributives are also distinct from restitutive/repetitive 'again', which relies on free or incorporating particles, as in *âuy-tsan* 'come back/again' or *poy/pegáu tsán* 'come again'.

(46) *Áwkau Ø- thón-dáw-de- em a- tsán- go- m.*
well 3SG-dug- be- BAS-LOC 1SG-arrive-DISTR-PFV
'I got around to places where wells had been dug.'
(Watkins 1984:234)

(47) *Thén Ø- gówbe-gu̠u- yi̠i.*
heart 3SG:3SG-miss- DISTR-IPFV.EVID
'He kept missing the heart [which was jumping about on the ground].'
(Harrington 1946:242)

The examples above therefore cover spatial distribution, whether of multiple achievements, looks, arrivals, or strikes. Event types can also necessitate a temporal distribution, as one participant cannot, for instance, visit multiple places at once.

Statives form their distributive with *-yáu*. They also are compatible with a single entity participating in multiple events (48).

(48) *Háundé tháymél ba- déy- yáu gau ba- kúu-yáu.*
what lonesomely 2PL-stand-DISTR and 2PL-sit- DISTR
'How lonesomely you are standing about and sitting about.'
(Watkins 1993:140)

This distributive also gives sortal readings, strewing events over a construed list rather than across space. In (49), the stative distributive indicates that each of the denotees of the implicit 'we' (only marked in agreement) has an age, which the modifier *hát* specifies as distinct from the others.

(49) *Hát ba- dá̠w-yá̠u.*
different 1IN.PL-be- DISTR
'We are different [ages].'

In (50), both distributives indicate that we have a set of animals, anaphoric in the context, and the speaker learns what each one does.

(50) *Háundé gyá- paul-go- m gau háyá*
something 3AN.PL:3PL-eat- DISTR-PFV and.SAME somehow
á- kíiyáu- de an yá̠- háaya- do.
3AN.PL-live.DISTR-BAS HAB :1SG:3PL-know.DETR.IPFV-because
'Because I find out what they [woodland animals] eat and how they live.'
(P. McKenzie 1987)

Distributive marking is not necessary for a distributive reading, as the following rendition of the Book of Genesis illustrates via an interesting pair. In (51a), *dáw* 'be'

bears a distributive, which *sául* 'be set' lacks, even though the trees are distributed about the garden. However, the reading of the distributive is sortal. Meanwhile, in (51b), it is *sául* that bears the distributive, with the spatial reading, while *dáw*, which could be distributive-marked, does not.

(51) a. *Áa Ø- sául t'áagya hát Ø- éytów-dáw-yáu dé- ey,*
tree 3SG-be set.PL pleasant different 3SG-fruit- be- DISTR BAS-LOC

  b. *áa Ø- sáw- yáu Ø- éytó-dáw-de.*
tree 3SG-be set.PL-DISTR 3SG-fruit-be- BAS
'Where there were trees of various pleasant fruits, fruiting trees all over.'
(Global Recording Networks n. d.: 1:391:48)

When distributives are used, the implicit number of locations and subevents is generally high, reflecting the notion of strewing, but the numerosity is approximative. This usage is distinct from nominal number, which is more precise about cardinality.

3 Agreement and the syntax of number

We begin our discussion of the syntax of number by recapitulating the key differences between agreement and suppletion touched on above and by highlighting morphological parallels between nominal and verbal marking, some of which may already have struck some readers. We then focus on agreement within the DP. Some DPs will host multiple inverse markers due to agreement, but many nominal modifiers, such as adjectives, numerals, and quantifiers are free of marking. Attributive uses of suppletive predicates which maintain number sensitivity are the sole number markers in the low DP. We advance the semantic generalisation that modifiers that take inverse marking concern particular individuals, a view that we support via morphological properties of incorporated nouns. Finally, we present a puzzling pattern of plural person agreement connected to indefinites.

Agreement and suppletion both track number but differ in two major respects. First, suppletion tracks actual cardinality.[8] Agreement proper often deviates from cardinality because it represents an amalgam of number and class. Deviations include the inverse, which can refer to singular, plural, or even, for first person exclusive, dual. Another deviation involves plural agreement for complex (nonatomic)

---

8 Exceptions occur but are rare. For example, when PPP nouns occur with plural-sensitive predicates, they only occur with the plural verb. When they occur with singular-sensitive stems, the verb reflects the participant number (Harbour 2008:141).

singular nouns (such as pluralia tantum nouns) and singular agreement for simplex plurals treated as atomic (such as collectives).

Second, suppletion is restricted to the innermost argument of the predicate (object or unaccusative subject), whereas agreement tracks the agent, applicative, and object, sometimes all at once. The three prefixes below illustrate this via an increase in morphemes that mirrors the increase in arguments, from intransitive *ma-* via transitive *mén-* to ditransitive *ménêy-* (which lowers the tone of the verb):

(52) *Ma- khîi!*
     2DU-exit.IMP
     'Come₂ out!'

     Mén-   khîi!
     2DU:3DU-exit.IMP
     'Take₂ them₂ out!'

     Ménêy-   khii!
     2DU:1/3.SG:3DU-exit.IMP
     'Take₂ them₂ out for me/him!'

The agreement system – one of the world's most intricate, compressing so much meaning into so little sound – cannot be treated adequately in this chapter. In relation to the themes of this volume, we note that person and number can fuse (e.g., :2DU: *m* is unrelated to :2SG: *g*), and that third persons and their numbers are visible to agreement (if they were not, the noun class system would vanish).

Amongst the numerous allomorphs of number/class agreement, one finds some reminiscent of nominal marking (Watkins 1984: 107–108). Many abstract nouns, like *kháwgya* 'name', *tǫwgyá* 'speech, word, language' and *k'ombáâlgya* 'imitation', involve the basic suffix *-gya*. These are PPP nouns and their suffix often resembles the plural agreement they govern, as in intransitive (53) (cf, 3AN.PL:3PL *gyá*, 3AN.PL:1/ 3.SG:3PL *gyâ*, 1SG:3PL *gyat*, (2/3.SG):1SG:3PL *yą́*):

(53) *Káuy- tǫ- **gya gya**-t'áagya.*
     Kiowa-say-<u>BAS</u> <u>3PL</u>- pleasant
     'Kiowa is pleasant.'

The rhymes of noun suffixes often coincide with agreement. For instance, nouns with basic forms ending in *-de* tend to be inherently paired, as are SDI *t'áwdé* 'ear' and *gúudé* 'horn', PPP *kháwdé* 'trousers', and SDP *p'ą́whį̂ide* 'half dollar'. The *e* vowel of *-de* occurs in many dual agreement prefixes, as in *mén-* 2DU:3DU and *ménêy-* 2DU:1/3.SG:3DU (52). Other similarities between rhymes of noun suffixes and agreement are evident in possessive agreement, both for inverse (54) and basic (55).

(54) *Pę́yg**aut** d**áut**-   dáw.*
    sand.<u>INV</u> :1<u>NSG</u>:<u>3INV</u>-be
    'The sand grain is ours.'

    *Pę̂y-   g**au** á**u**-   dáw.*
    turkey-<u>INV</u> :3<u>SG</u>:<u>3INV</u>-be
    'The turkeys are his/hers.'

(55) *Khíi-d**á** **á**-   dáw.*
    day-<u>BAS</u> :3<u>SG</u>:<u>3SG</u>-be
    'The day is his/hers.'

    *Táa-d**e** n**é**-   dáw.*
    eye-<u>BAS</u> :1<u>SG</u>:<u>3DU</u>-be
    'The eyes are mine.'

The correspondences are historic rather than synchronic: as many examples in this chapter show, agreement and noun suffix often diverge in phonological form. Moreover, the corresponding segments above have taken on broader functions in both nouns and verbs: there are paired PPP nouns in *de*, like *kháwdé* 'trousers', and non-inverse prefixes with *au*, like *gau* 3:2SG. There are, furthermore, exponents of number confined to a single domain, which we would not expect if the actual forms of the noun and agreement were directly linked. For instance, the nasalization that frequently occurs with dual agreement, as in (52) and (55), is not a feature of number on nouns. These correspondences suggest that various suffixes on nouns are vestiges of singular, dual, and plural morphemes. Synchronically, nouns only contrast inverse versus basic forms and express semantic number only through an amalgam of noun and class.

    Complex DPs can host multiple occurrences of inverse or basic marking, as when demonstratives modify nouns (13a)–(13b). Other complex DPs that contain multiple number marking include nouns with the quantifier 'only' (56) and possessed nouns with anaphoric third person possessors (57). In both of these, multiple inverse markers can even be adjacent.[9] For anaphoric third person possession, a prefix marks possessor person and a suffix, BAS/INV. The noun can additionally bear inverse marking and both instances of inverse marking can reflect possessum number (as in 'his/her children').

(56) *táttau-   gau-ki*
    skunk.<u>INV</u>-<u>INV</u>-only
    'only skunks'

---

9 Inverse agreement is optional for 'only', permitting *táttau-de-ki* [skunk.INV-BAS-only] as an alternative to (56). This is the only case of optional nominal inverse in the language.

(57) *á- ii- te*
3POSS-child-BAS
'her/his child(ren₂)'

á- yyoy- gau
3POSS-child.INV-INV
'her/his/their children₃₊'

These multiple occurrences notwithstanding, there are comparatively few loci of number marking within DP. Adjectival predicates have none. Instead, they form a complex word with the noun and the complex as a whole is marked. For instance, 'white horse(s₂)' can be either head-final *t'ạ́y-tsệy* or head-initial *tsệy-t'ạy*. The inverse suffixes to each of these, its form conditioned by its neighbour: *t'ạ́ytsệygau*, *tsệyt'ạymau* 'white horses₃₊'. A rare three-way number distinction is available in SDI and IDI nouns modified by a suppletive adjective of size (cf, table 4). The inverse, again, appears on singular and plural forms, conditioned phonologically by the adjective, but, additionally, the adjective itself suppletes for nonsingular. This combination of factors leads to distinct singular, dual, and plural forms.

(58) *ául- kyọ́y- máu*
hair-long.SG-INV
'a long hair'

ául- kịiníi
hair-long.NSG
'long hairs₂'

ául- kịinóp
hair-long.NSG.INV
'long hairs₃₊'

Despite the complexity of morphological number, numerals and fractions in Kiowa are very simple. They neither take number marking nor do they affect marking on nouns and verbs. For instance, singular nouns of any class use invariant *páagau* 'one' and maintain whatever agreement noun class demands, whether singular, plural, or inverse (59).[10]

(59) *Páagau {pól Ø- dáw} / {kút gya- dáw} / {tóp e- dáw}.*
one bug 3SG-be book 3INV-be peg.INV 3INV-be
'It is one bug / book / peg.'

---

10 *Páagau* 'one' and *páa* 'some' are likely historically related, derived via a suffix related to *-kaw* 'just, only, and no more', used with other numbers and quantifiers (*yíikaw* 'just two', *háotekáw* 'just a few'; hence, the use of *páagau* to mean 'lone', as in *Kûypaagauy* 'Lone Wolf').

Similar facts hold for (60) 'two', (61) 'three', and other numerals. No verb agreement options other than those shown are permitted:

(60) *Yíi k'yátáy-k'ïi / *k'yátây e̱- dáwmêy.*
    two chief- MALE chief.INV 3DL-be.EVID
    'There were two chiefs.'

(61) *Phá̱aow sá̱adau / *sân é- tsán.*
    three child.INV child.BAS 3INV-arrive.PFV
    'Three children arrived.'

Matters are the same for 'half', the only fraction we have recorded. For half of a single object, the agreement triggered is that of the whole (singular for 'horse', inverse for 'stick'):

(62) *Tsê̱y zâyde gya- bó̱w.*
    horse half 1SG:3SG-see.PFV
    'I saw (one side of) the horse.'

    Áa- dau zâyde dé- bó̱w.
    stick-INV half 1SG:3INV-see.PFV
    'I saw (half) the stick.'
    (based on Harrington (1928:203))

Our few examples of integers plus fractions, like 'one and a half cookies', are expressed via disjoint conjunction and so conform to the description above:

(63) *Páagau éyk'audal gya- hân gau zâyde-al.*
    one cookie 1SG:3SG-eat up.PFV and half- also
    'I ate one and a half cookies.'

Numerals, like other nominal modifiers, are syntactically flexible in Kiowa, preceding or following the noun, sometimes nonadjacently (Adger, Watkins, and Harbour 2009) In (64), *yáukáuy* 'young woman, young women₂' can occur in any of the three positions shown. In all cases, the noun must be unmarked and the verb, dual.[11]

---

11 Numerals are always nominal modifiers and cannot be inflected like verbs (below left). If incorporated into the verb 'be', they produce an ordinal reading, rather than 'are five' (middle). The same ordinal reading is available in nominal compounds (right).

(i)   *E- áunt'aw.*   *E- áunt'au-daw.*   *yíikyá-phii-gau- em*
    3INV-five   3INV-five- be   four- hill- NOM-LOC
                      'They are fifth.'   'at the fourth hill'
                      *'They are five (in number).'

(64) (*Yáukáuy*)   *yíi* (*yáukáuy*)   *nen-* *bǫ́w*   (*yáukáuy*).
young woman two young woman 1SG:3DU-see.PFV young woman
'I saw two (young women).'

Quantifiers too lack inverse marking. Like 'half', 'some' occurs with the verb agreement form dictated by its head noun, inverse in (65) but plural in (66). This holds even if the head noun (in parentheses) is omitted.

(65) *Kâul*   (*áadauattau*) *có* *dé-*   *kâul-*   *taw.*
some.IRR barrel.INV   thus 1SG:3INV-turn over-MOD
'I could turn one (barrel) over.' (P. McKenzie n. d.-b)

(66) *Kâul*   (*hǫ́wsauaa*) *gyat-* *áum-* *táw.*
some.IRR fence posts 1SG:3PL-make-MOD
'I am going to make some (fence posts).'
(Goins 1957)

This is also true of *téy* 'all', as inverse (67), plural (68), and singular nouns (69) illustrate.

(67) *Téy* (*k'ǫ́weytǫw-baut*) *náu-*   *ǫ̨w.*
all frybread- INV   2SG:1SG:3INV-give.IMP
'Give me all (the piece of frybread).'
(Ahote 1965)

(68) *Téy gyá-*   *thaa-*   *hel* ... *á-*   *kíi-* *ǫa-* *hel-* *gau.*
all 3AN.PL:3PL-cut.PL.OBJ-EVID   3AN.PL-meat-come-EVID-INV
'They who came for meat cut all [the meat] away.'
(Toyebo 1957a)

(69) *Téy an*   (*áulhǫ́w-gya*) *á-*   *hauttau.*
all HAB money- BAS 3AN.PL:3SG-take.IPFV
'They take it all (the money).'
(Redbird 1957)

Wh-elements display mixed behaviour. Some, like *háote* 'how many', are invariant, whether their head noun is inverse marked (70), basic (71), or absent (parentheses).

(70) *Háote*   (*sǫ̨adau*) *gáu-* *dáw?*
how many child.INV :2SG:3INV-be
'How many (children) do you have?'
(Watkins 1984:214); *gáu* for *gáut*)

(71) *Háote   (áa) gyá-   t'au- thátkyá?*
how many tree :2SG:3SG-chop-sever.PL.DETR.PFV
'How many (trees) have you gotten chopped up?' ('tree' is IDS)

Others must be basic or inverse in accord with their head noun, like 'which' with singulars of SDS *máun* 'finger' and IDP *máun* 'hand' (72)–(73) (also (18)–(19)).

(72) *Hâagyây   máun gyá-   kówlí- dáw?*
which.BAS finger :2SG:3SG-numb-be
'Which of your fingers is numb?'

(73) *Hâabâw   mą́wdáu é-   thêm?*
which.INV arm.INV 3SG:3INV-break.PFV
'Which arm did he break?'

The modifiers that take inverse marking are, apparently, those with a definite individual built in. Thus, demonstratives, possessives, and the *wh*-elements *hâagyây/hâagâw* 'which.BAS/INV' and *hâundé/hâungáu* 'what, who.BAS/INV' do (cf. obligatory marking on relative clauses (30)–(31)). Numerals, fractions, quantifiers like 'all' and 'some', and *wh*-elements like 'how many' do not. The fit is not perfect, however: 'only' is optionally marked, *(ts')al* 'too' never is; *hâundé/hâungáu* 'what, who' distinguish inverse/basic but *hâatêl* 'who' does not (it is restricted to singular/dual; Watkins, p.c.).

A link between inverse marking and definite individuals suggests that the inverse belongs, in theoretical terms, to the higher reaches of the DP. The absence of inverse and basic marking on incorporated nouns supports this view. Although Kiowa generally lacks direct object incorporation, incorporation for other purposes is highly productive. The incorporated noun is number-neutral (Watkins 1984; Adger, Harbour, and Watkins 2009; A. McKenzie 2017, 2019, 2021). Representative examples, alongside (3), are (74), where bare SDI 'horse' can have any number reference, (75), where SDI 'neck' has plural reference but is bare *k'ól*, not inverse-marked *k'ôwtau*, and (76), where *kyákôm-da* 'life-BAS' occurs without its basic marker:

(74) *a- tsêy(*gau)- to̱-   baa*
1S-horse(INV)-look for-go.PFV
'I am going to look for a horse/horses.'

(75) *Béthaw   ét-   k'ól/*k'ôwtau- thaa-   hel.*
apparently 3INV:3INV-neck/neck.INV-sever.PL.OBJ-EVID
'Apparently they had cut through their throats.'
(P. McKenzie n. d.-c)

(76) *P'áw- tháw- be   á- k'yákôm(*da)-t'aw.*
water-beyond-along 3PL-life(BAS)-   stay
'Across the ocean there are people living.'
(Crowell 1960)

Incorporated nouns are typically taken to be smaller than full DPs. By associating number marking with the higher functional domain, we correctly derive its absence under incorporation.

We conclude this section with a fascinating construction. When the indefinite *háagyây* 'which(ever)' occurs without a head noun, agreement on the verb corresponds to the group over which it ranges.[12] In (77), the meaning 'whichever of us' emerges from *háagyây* 'which(ever)' plus first inclusive nonsingular agreement (*ba-*). Strikingly, the nonsingular agreement extends into the second clause (*bét-*), which ought to be singular as it describes what the winner will do. Descriptively, it is as though 'whichever' (or the noun phrase it belongs to) escapes its conjunct to scope over the whole sentence.

(77) *Háagyây ba-   t'aum-tsán- t'aw gigáu   bét-   hân.*
which.BAS 1IN.NSG-first- arrive-MOD and then.SAME 1IN.NSG:3INV-eat up.PFV
'Whichever of us comes first will eat them up.'
(Harrington 1946:239)

Plausibly, (77) involves "unagreement," first or second person agreement paired with a common noun that lacks any indication of person (Hurtado 1985). Another example is (29) where the meaning 'us old men' arises from *êlk'yoy* 'old men' with 1NSG agreement. A semi-technical gloss of (77) would be 'us-whichever [i.e., some one of us] will come first and then us-that [i.e., that one of us] will eat them up', which involves an indefinite in one clause, null resumption in the next, and unagreement in both. The details of this approach lie beyond the scope of this chapter (see Harbour 2022). We are aware of analogous constructions only in Georgian (Léa Nash, p.c.) and, more transparently, in Huallaga (Huánuco) Quechua (Weber 1989:308).

4 Semantics and discourse

At the level of discourse semantics, Kiowa number is rather uncomplicated. As already discussed (72)–(73), inverse marking, or its absence, on interrogatives can indicate what answer a speaker anticipates. Other discourse-semantic properties not illustrated so far include: that nouns with plural reference can refer to singulars

---

12 Interrogatives with initial falling tone (e.g., *hâagyây* 'which', *hâundé* 'what, who') have matching indefinites with initial high (*háagyây* 'which(ever)', *háundé* 'something, someone').

under some circumstances; that number has implications for politeness (for third persons, with dual sometimes emerging as the culturally apt option); and that forms capable of plural reference are used in generic statements (consistent with plural being the default number; section 2.3). Although number is involved both in agreement and switch reference, it is not a driving force of their pragmatics but is secondary to other discourse constraints. In these regards, Kiowa is not especially unfamiliar. Discourse structure does license one surprising phenomenon in the domain of number though, namely, "associative conjunctions," which are encoded via a single name plus nonsingular agreement. In our examples, these require familiarity of the conjuncts and greater discourse prominence for the named person. We address these subjects in the order just given.

Nouns with plural reference permit nonplural reference under some circumstances. For instance, (78), with inverse *thaalyóp* 'boys$_{3+}$', is false if even one boy arrived.

(78) *Háun kâul thaalyóp e-   tsą́anâw.*
     NEG   some boy.INV 3INV-arrive.NEG
     'No boys arrived.'

Similarly, the instruction in (79) about collecting honey has inverse subject agreement, referring to previously mentioned (semantically plural) bees. Yet, this advice presumably applies even if a noninverse number of bees (one or two) land on your face:

(79) *T'ów-ba   dét-   thaw- gún- taw- al   poy   háyá*
     face- against 3INV:2SG:RX-sit.NV-jump-MOD-also PROH somehow
     em- thǫw- kówbį̀i-   taw.
     2SG-chase-thrash about-MOD
     'Even if they land on your face, don't thrash about.'
     (P. McKenzie 1987)

Plurality is not used for politeness, nor is dual used for intimacy. However, there is a special form of agreement used for animates subject to empathy, generally other Kiowas but not members of other groups – though non-Kiowa (81), or even nonhuman (50), referents are possible if the setting is right. This animate plural agreement (AN.PL) overrides inverse agreement; compare (80a) with (80b) or the two verbs in (81).

(80) a. *Káuy- gú á-   hóuąa- hel*
        Kiowa-INV 3AN.PL-travel.PFV-EVID
        'The Kiowas were travelling along'

b. *Kyây- gu e- hóuah- hel*
Comanche-INV 3INV-travel.PFV-EVID
'The Comanches were travelling along'

(81) *Kíi táa- hey an gyá- pauttau máun kâul e- áun.*
meat cook-PRIV HAB 3AN.PL:3PL-eat.IPFV maybe some 3INV-think
'Some [of the white people] maybe thought they [Kiowas] eat meat raw.
(P. McKenzie n. d.-f)

An interesting complication is that animate plural agreement does not exist for se-
mantically plural internal arguments (objects, themes) in the presence of an appli-
cative. These cooccur, for instance, in possessives, with the applicative agreement
encoding the possessor of the internal argument, as the contrast between one-argu-
ment, nonpossessive (82a) and two-argument, possessive (82b) illustrates:

(82) a. *Yíikya hegáu sḁadau e- dáw.*
four then child.INV 3INV-be
'There were four children then.'

b. *Yíikya hegáu sḁadau dáut- dáw.*
four then child.INV 1NSG:3INV-be
'We had four children then.'
(P. McKenzie 1986)

The possessum of (82b) is semantically an animate plural and it is expressed by
inverse number. Inverse number for referents that would normally command em-
pathic AN.PL agreement can be felt to be culturally inapt. Speakers have recourse
to two strategies, depending on the person of the applicative.

First, dual may be used instead of inverse. In (83), inverse *maayóp* 'women'
cooccurs with such dual agreement, even though no Kiowa noun is inverse-marked
when semantically dual:

(83) *Maayóp nén- háygyá-daw.*
woman.INV :2SG:3DU-known- be
'You know women.'
(Watkins 1984: 146)

A direct comparison of the acceptability between inverse and dual in this context is
provided by (84):

(84) *Káuy- gú né- tsán / (*)náu- tsán.*
Kiowa-INV :1SG:3DU-arrive.PFV :1SG:3INV-arrive.PFV
'The Kiowas came to me.'
(Harbour 2008:72)

For third person applicatives, a second strategy is available: reflexive agreement, with the animate plural marked as an agent, even if the verb is intransitive.

(85) *Tsólhautkau Dawk'íi ém-　　　áwdéy-dą̄wmêy.*
　　　thus.INV　　God　　3AN.PL:RX-dear-　be.EVID
　　　'Such people are dear to God.'
　　　(Toyebo 1957a)

Generic statements often take animate agreement for human generic subjects. For instance, (86) can be used to silence a child making a fuss at not getting something:

(86) *Háun an　á-　　hę̄ymâw!*
　　　NEG　HAB　3AN.PL-die.NEG
　　　'They don't die [of such things].'

The literal reading 'They don't die' is more readily rendered in English by an impersonal or second person 'One doesn't *or* you don't die [of such things]'. A similar example of Kiowa using animate plural where English would use 'you' or 'one' is the following statement about travel preparation:

(87) *Kólbél　gya-　　máwkhól-　daw háyá*
　　　securely　:3AN.PL:3PL-prepare.NV-be　　somewhere
　　　á-　　hów-　banma- tsey.
　　　3AN.PL-travel-go.IPFV-when.SAME
　　　'One is to be [lit. they are] well prepared when [lit. they are] going to travel somewhere.'
　　　(P. McKenzie n. d.-g)

Plural is the default in other kinds of generic sentences, in keeping with plural as the default number more broadly (32)–(35). For instance, a generic statement about coyotes (SDI) and their noses (IDP) requires 'coyote' in the inverse and 'nose' in the (basic) plural, as (88b) illustrates. In contrast, the episodic statement in (88a) about Coyote and his nose require the singular and inverse respectively.

(88) a. *Áwgau Séndé mawthǫtsówhį̄i mawk'âun áu-　　　sǭw- hêl*
　　　　REL　Sende coyote.BAS　　nose.INV　3SG:3SG:3INV-hone-EVID
　　　　de- peydow,
　　　　BAS-because

　　　b. *ę́yhaw- al　hétáu mawk'áun máw tsǭw-gáu bét-　　dáw.*
　　　　now-　also still　nose.BAS　like　awl- INV　:3INV:3PL-be
　　　　'Because Sende honed the Coyote's nose, still nowadays their noses are like awls.'
　　　　(Harrington 1946:240)

Agreement and switch-reference are central to tracking discourse referents in Kiowa (Watkins 1993; A. McKenzie 2012, 2015). Number is implicated in both these systems, but does not itself play a major role in them. Consider agreement. In *Sende and the Mountain Ogres* (Harrington 1946:240–242) it is hard to find a run of three sentences together where neither Sende nor the ogres are mentioned by noun phrase. One might have expected singular agreement for Sende and inverse for the ogres to be enough to identify who does what to whom. Evidently, though, number alone is felt to be pragmatically or stylistically insufficient.

Switch-reference marking is often triggered between two sentences when the first identifies a group and the second picks out a subset. In (89), the first clause has an animate plural subject, while the second has third singular, corresponding to one member of the plurality. The two are linked by switch-reference marker *-nau* (as opposed to the second and third clauses, which are linked by a same-reference marker, *gau*).

(89) *Á-   maw- taw- nau   k'yá̱ahî̱i tsê̱y- gau*
     3AN.PL:3SG-move-MOD-when.DIFF man   horse-INV
     é-   pawkâun-   taw gau   hágyá
     3SG:3INV-bring along-MOD and.SAME somewhere
     em-   áw-   saw-taw.
     3SG:RX-temporarily-sit- MOD
     'When they moved camp, the man would bring the horses and sit awhile.'
     (Redbird 1957)

However, number and subsets are not the driving factor here. In Kiowa, switch-reference across coordinated clauses is "non-canonical" in Kiowa: It does not express the identity or disjointness of the subjects (Haiman and Munro 1983). If the speaker envisages the two conjuncts as constituting a single situation, then same-referent connectives are used (A. McKenzie 2011, 2012), whether with referential (90) or quantificational (91) subjects.

(90) *Kathryn gya-   gút   gau   Esther-al gya-   gút.*
     Kathryn 3SG:3PL-write.PFV and.SAME Esther-also 3SG:3PL-write.PFV
     'Kathryn wrote a letter and Esther wrote one too [*e.g.*, in a campaign].'

(91) *Étté   thó̱wtsép hágyá á-   kawley   gau   páa*
     many flood   once 3AN.PL:3SG-cross.IPFV.EVID and.SAME some
     á-   ó̱ba- hi̱i- hel.
     3AN.PL-drown-die-EVID
     'Many were crossing a flood once and some drowned.'
     (Watkins 1984:159)

The discourse factors above affect switch reference only with coordinating connectives. With subordinating connectives *tsey* 'when.SAME' and *ey* 'when.DIFF', subject number becomes a factor. In (92), the same-reference marker is ungrammatical.

(92) *Háatêl   Ø- tsán-   ey,   téy ém-   kún-   haa.*
     someone 3SG-arrive.PFV-when.DIFF all 3AN.PL:RX-dance.NV-get up.PFV
     'When someone [specific] showed up, everyone got up to dance.'
     (A. McKenzie 2012:209)

A more truly discourse-driven use of number involves implicit or associative conjunctions. These involve nonsingular agreement for a personal name or kinship term that would ordinarily govern singular agreement. In (93) and (94), for example, only Satanta and Adam are named, but by dual agreement in both signals references to pairs of individuals, Satanta and Big Tree, and Adam and Eve respectively:

(93) *Sétt'ą́ydé hegáu óópkau Tȩhą́ane-ku et-   ą̀a- hii- hel.*
     Satanta   then   far away Texas-   to 3INV:3DU-haul-move-EVID
     'They carted Satanta [and Big Tree] far away to Texas.'
     (P McKenzie n. d.-a)

(94) *Édam ȩ- khóbéttáu-dáw dé- tso náw-ál   páatsokáw*
     Adam 3DU-sin-   be BAS-thus 1-   also likewise
     ba-   khóbéttáu-dáw.
     1IN.NSG-sin-   be
     'We are sinful as Adam [and Eve] were.'
     (Global Recording Networks n. d.: 5:325:41)

Watkins (p.c.) observes that, in the texts from which (93)–(94) come, Big Tree and Eve have already been in conjunction with Satanta and Adam and the mentioned member of each pair has greater prominence. For Satanta, this relates to his not surviving the incident that the text recounts. For Adam, it relates to his being present and more active for most of the story. That parts of conjunctions are dissociable was already seen in the split conjunction in (17). More relevant here are examples like (95), where a conjunction of two singulars splits across a verb with dual agreement:

(95) *Ḗygau Dawk'yaíi Ø- dáw-dé- taul ȩ- dáwmêy   gau á-   tsaw- de.*
     here Jesus   3SG-be- BAS-father 3DU-be.IPFV.EVID and 3POSS-mother-BAS
     'Here were Christ's father and his mother.'
     (P. McKenzie n. d.-h)

Null anaphora in lieu of the second conjunct may underlie associative, or implicit, conjunctions. Null arguments are licensed under similar conditions of discourse familiarity (Watkins 1990).

5 Conclusions

A morpheme like the Kiowa inverse, which makes some nouns singular, others plural, and others both, is sure to capture the attention of typologists and theoreticians alike. However, the inverse is the just the start of the grammar of number in Kiowa. In a rare form of linguistic economy, different distributions of that one suffix serve to define four noun classes. Moreover, the system of inverse marking is embedded within a singular-dual-plural number system, and noncanonical uses of these numbers – such as singular for plural and *vice versa* – reveal another four noun classes. Again, this is a radical economy, using independently available numbers as the means of defining noun classes. These properties make for a language in which number and class are inextricably linked and where class is expressed with remarkable morphological economy.

Despite the minimality of these morphological resources, Kiowa is morphologically highly intricate, especially as concerns the tracking of number by the verb. The agreement prefix encodes singular, dual, plural, and animate plural agreement, plus all information about class, whether inverse, singular-for-plural, dual-for-plural, or plural-for-nonsingular, and does so while registering up to three arguments in as many as four persons. These complications mean that number agreement diverges from true cardinality. The most reliable expression of cardinality occurs in the most irregular corner of the language, in the suppletive system. Jointly, agreement and suppletion completely encode number and class. So, the verb is overwhelmingly the morphological locus of core nominal information.

The interest of Kiowa number does not end there. Some properties of the system may be straightforward, such as the inertness of numerals and the availability of singulars from forms that refer to plurals. However, other constructions are more unusual: agreement in *whichever* relatives, implicit or associative conjunctions, and the emergence of dual as the culturally appropriate means of reference for animate plurals.

Kiowa is the best documented member of the endangered Kiowa-Tanoan family, yet it is something of an outlier morphologically. The Tanoan languages all inverse-mark (eligible) nouns in the dual (e.g., Tewa, Harrington 1910a; Taos, Harrington 1910b; Jemez, Yumitani 1984) and some (e.g., Tewa, Jemez, but not Taos) mark personal pronouns in the dual and plural. The extent of singular-for-plural and plural-for-singular agreement is less well documented. And many of their agreement systems are less well articulated, with dual and inverse frequently collapsing (e.g., in Jemez). Given these variations on such a rare design of number system, it would be fascinating to address the questions that underlie the current volume for the family as a whole.

Abbreviations

1	first person
2	second person
3	third person
ADV	adverb
AN	animate
AUX	auxiliary
BAS	basic
DETR	detransitive
DIFF	different subject
DISTR	distributive
DU	dual
EVID	evidential
EX	exclusive
HAB	habitual
IMP	imperative
IN	inclusive
INDEF	indefinite
INV	inverse
IPFV	imperfective
IRR	irrealis
LOC	locative
MALE	male
MOD	modal
NPL	nonplural
NSG	nonsingular
NEG	negative
NV	nonverbal
OBJ	object
PFV	perfective
PL	plural
PRIV	privative
PROH	prohibitive
PROX	proximative
REL	relative
RX	reflexive
SAME	same subject
SG	singular

References

Adger, David, Daniel Harbour & Laurel J. Watkins. 2009. *Mirrors and microparameters: Phrase structure beyond free word order*. Cambridge: Cambridge University Press.

Ahote, Alma. 1965. Kiowas in Amarillo. Summer Institute of Linguistics Recording.

Comrie, Bernard. 2001. How independent is Tsezic morphology? In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *CLS 37: The panels: 2001: Proceedings from the*

parasessions of the 37th Meeting of the Chicago Linguistic Society, 367–383. Chicago, IL: Chicago Linguistics Society.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Crowell, Edith. 1960. *The Kiowa language: A grammatical study*. Philadelphia, PA: University of Pennsylvania dissertation.

Doron, Edit & Ana Müller. 2014. The cognitive basis of the mass-count distinction: Evidence from bare nouns. In Patricia Cabredo Hofherr & Anne Zribi-Hertz (eds.), *Crosslinguistic studies on noun phrase structure and reference* (Syntax and Semantics 39), 73–101. Leiden: Brill.

Global Recording Networks. n. d. Words of life: Kiowa. http://globalrecordings.net/en/language/kio (accessed 14 February 2021).

Goins, Lucy. 1957. Gardening. Summer Institute of Linguistics Recording.

Haiman, John & Pamela Munro (eds.). 1983. *Switch-Reference and Universal Grammar*. Amsterdam: John Benjamins.

Harbour, Daniel. 2008. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.

Harbour, Daniel. 2022. Plural agreement for singular *whichever* of: A Kiowa-Kartvelian conundrum. In Nora Boneh, Daniel Harbour, Ora Matushansky, & Isabelle Roy (eds.), *Building on Babel's rubble*, 411–420. Paris: Presses Universitaires de Vincennes.

Harbour, Daniel, Laurel Watkins & David Adger. 2012. Information structure, discourse structure, and noun phrase position in Kiowa. *International Journal of American Linguistics* 78(1). 97–126.

Harrington, John P. 1910a. A brief description of the Tewa language. *American Anthropologist* 12. 497–504.

Harrington, John P. 1910b. An introductory paper on the Tiwa language, dialect of Taos, New Mexico. *American Anthropologist* 12. 11–48.

Harrington, John P. 1928. *Vocabulary of the Kiowa language* (Bulletin 84). Washington, DC: Bureau of American Ethnology.

Harrington, John P. 1946. Three Kiowa texts. *International Journal of American Linguistics* 12(4). 237–242.

Hurtado, Alfredo. 1985. The unagreement hypothesis. In Larry D. King & Catherine A. Maley (eds). *Selected papers from the thirteenth Linguistic Symposium on Romance Languages*, 187–211. Amsterdam, the Netherlands: John Benjamins.

Lima, Suzi. 2014. *The grammar of individuation and counting*. Amherst, MA: University of Massachusetts Amherst dissertation.

McKenzie, Andrew. 2011. Subject domain restriction and reference-tracking. In Nan Li & David Lutz (eds.), *Proceedings from the 20th conference on Semantics and Linguistic Theory (SALT 20)*, 269–288. LSA eLanguage. https://journals.linguisticsociety.org/pro (accessed 14 February 2021).

McKenzie, Andrew. 2012. *The role of contextual restriction in reference-tracking*. Amherst, MA.: University of Massachusetts Amherst dissertation.

McKenzie, Andrew. 2015. A survey of switch-reference in North America. *International Journal of American Linguistics* 81(3). 409–448.

McKenzie, Andrew. 2017. Sources of intensionality in Kiowa noun incorporation and English synthetic compounds. Colloquium. University of Massachusetts Amherst.

McKenzie, Andrew. 2019. The role of semantics in licensing English synthetic compounds. In Richard Stockwell, Maura OLeary, Zhongshi Xu, & Z. L. Zhou (eds). *Proceedings of the 36th West Coast Conference on Formal Linguistics (WCCFL 36)*, 281–285. Somerville, MA: Cascadilla Proceedings Project.

McKenzie, Andrew. 2020. Two types of habituals: Kiowa ingredients of a modular imperfective. Paper presented at Sinn und Bedeutung 25. University College London & Queen Mary University of London, 2 September.

McKenzie, Andrew. 2021. Mediating functions and the semantics of noun incorporation. *Natural Language and Linguistic Theory*. DOI: 10.1007/s11049-021-09511-4.

McKenzie, Parker. 1963a. Letter to Charlie Redbird. October 15. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. 1963b. Letter to Charlie Redbird. February 16. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. 1986. Black widow, or tornado. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. 1987. An encounter with honey bees. Re-telling of a story by Bert Geikaumah. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. 1990. Letter to Laurel J. Watkins. April 29. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. 1993. Náu: fèljàugà [My thoughts]. Manuscript, Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. n. d.-a. The time they imprisoned Kiowa chiefs. Ms. transcription of telling by Bert Geikaumah. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. n. d.-b. Jack Wolf and the barrel. Recorded and glossed by Laurel Watkins.

McKenzie, Parker. n. d.-c. Cutthroat massacre. Recorded by Laurel Watkins. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. n. d.-d. Runaway horses. Recorded and glossed by Laurel Watkins.

McKenzie, Parker. n. d.-e. Grandfathers calf. Recorded and glossed by Laurel Watkins.

McKenzie, Parker. n. d.-f. Events at a Kiowa beef butchering. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. n. d.-g. Written examples, part 2. Ms. Oklahoma Historical Society Research Center, Oklahoma City, OK.

McKenzie, Parker. n. d.-h. Christmas encampment. Re-telling of a story by Linn Pauahty. Recorded and glossed by Laurel Watkins. Oklahoma Historical Society Research Center, Oklahoma City, OK.

Redbird, Charles. 1957. Old times. Summer Institute of Linguistics Recording.

Redbird, Charles, Carrie Redbird, William Wolfe, Louis Toyebo & Lorna Gibson. 1962. *Kawydawkhyah dawgyah* [Kiowa Christian songs]. Norman, OK: SIL International.

Spotted Horse, Helen. 1957. Sende meets a Whiteman. Summer Institute of Linguistics Recording.

Toyebo, Louis. 1957a. A kind man. Summer Institute of Linguistics Recording.

Toyebo, Louis. 1957b. Boy hunting. Summer Institute of Linguistics Recording.

Toyebo, Louis. 1957c. Jonah and the whale. Summer Institute of Linguistics Recording.

Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln, NE: University of Nebraska Press.

Watkins, Laurel J. 1990. Noun phrase versus zero in Kiowa discourse. *International Journal of American Linguistics* 56(3). 410–426.

Watkins, Laurel J. 1993. The discourse function of Kiowa switch-reference. *International Journal of American Linguistics* 59(2). 137–164.

Watkins, Laurel J. & Daniel Harbour. 2010. The linguistic genius of Parker McKenzies Kiowa alphabet. *International Journal of American Linguistics* 76(3). 309–333.

Weber, David. 1989. *A grammar of Huallaga (Huánuco) Quechua* (UC Publications in Linguistics 112). Berkeley, CA: University of California Press.

Wolfe, William. 1957. Happening in a corn field. Summer Institute of Linguistics Recording.

Wonderly, William, Lornia Gibson & Paul Kirk. 1954. Number in Kiowa: Nouns, demonstratives, and adjectives. *International Journal of American Linguistics* 20. 17.

Yumitani, Yukihiro. 1984. *Case marking in Kiowa: A study of organization of meaning*. New York, NY: City University of New York dissertation.

Marianne Mithun
20 Number in Mohawk (Iroquoian)

Abstract: Mohawk is a prototypical polysynthetic language: verbs are holophrastic, capable of serving as complete sentences in themselves. Interestingly, number distinctions are centered in verbs rather than nouns. All verbs are finite, with obligatory pronominal specification of their core arguments. The pronominal prefixes distinguish first and second persons, and neuter, masculine, feminine-zoic, and third persons. There are grammatical agent, patient, and transitive paradigms. Singular, dual, and plural number are distinguished for first and second persons and masculine and feminine-zoic for third person agents. Only singular and duo-plural are distinguished for masculine and feminine-zoic patients. Number is not generally distinguished for neuters (inanimates). Nouns, by contrast, are grammatically neuter, so do not distinguish number. (Nominals referring to persons are morphological verbs.) Inalienable possessive prefixes on nouns mirror the agent forms on verbs, and alienable forms those of patients. They do not, however distinguish number or show agreement with the pronominal prefixes on verbs.

1 Overview

Mohawk is a language of the Iroquoian family, indigenous to northeastern North America. There are six main Mohawk communities, located in Quebec, New York State, and Ontario. The language is still spoken skillfully by elder first-language speakers, and now fluently by ever-growing groups of second-language speakers.

Typologically, Mohawk is polysynthetic under any definition of the term. Iroquoianists distinguish three lexical categories, defined according to their morphological structure: verbs, nouns, and particles. Verbs, which can be morphologically quite complex, contain obligatory inflectional pronominal prefixes identifying core arguments. The grammatical relations distinguished are not the more familiar subjects and objects, or ergatives and absolutives. Instead they are grammatical agents and patients. Agents are normally those who are in control, while patients are those not in control but affected, including recipients and beneficiaries. Speakers do not make online semantic distinctions as they speak, however: the pronominal paradigm is lexicalized with each verb. There are three sets of pronominal prefixes: one for grammatical agents, one for grammatical patients, and one for transitive combinations. There is noun incorporation, a construction in which a noun stem is combined with a verb stem to form a larger verb stem. Aside from imperatives, all verbs are obligatorily inflected for aspect. Verbs may also contain a variety of additional prefixes and/or suffixes, in a mainly templatic structure. Basic nouns show simpler morphological structure, with just a gender or possessive prefix and a noun suffix

which adds no further meaning. Particles are by definition morphologically simplex, though they may be compounded.

These lexical categories are not, however, always isomorphic with syntactic function. In particular, morphological verbs can function as predicates, as in other languages, but also as complete sentences in themselves, and as referring expressions, with no additional marking. Morphological nouns are used only as referring expressions. Morphological particles, which by definition are words with no internal morphological structure, serve a wide variety of syntactic and discourse functions. Among the particles are some independent pronouns used only in certain contexts, demonstratives, and indefinite and interrogative pronouns.

Gender is distinguished in the third person pronominal prefixes on verbs, in the possessive prefixes on nouns, and in third person particle pronouns. There are four genders: masculine, feminine-zoic, feminine-indefinite, and neuter. Masculine forms are used for male persons and a few animals. Feminine-zoics are used for some female persons and most animals. Feminine-indefinites are used for other female persons, those whose sex is unspecified, and generic persons. Neuters are generally used for inanimate objects. Demonstratives, indefinite pronouns ('someone', 'something', 'anyone', 'anything', and interrogative pronouns) do not distinguish gender.

There are singular, dual, and plural number categories, but not all numbers are distinguished in all types of words or in all their forms. Three numbers are distinguished on first and second person pronominal prefixes on verbs and possessive prefixes on nouns, but none on first and second person particle pronouns. There are only singular forms for feminine-indefinite pronominal prefixes, and number is not distinguished for neuters. Singular, dual, and plural number are distinguished in agent prefixes on verbs and inalienable possessive prefixes on nouns, but only singular and duo-plural are distinguished in patient prefixes on verbs, alienable possessive prefixes on nouns, and third person particle pronouns. The number distinctions that are available are obligatory. Demonstratives, indefinite pronouns, and interrogative pronouns do not distinguish number.

Verbs may also optionally contain distributive suffixes, which distribute events and states over various locations, times, or participants. Nouns may contain distributive enclitics, which distribute types over individuals: 'various and assorted'. Both sets of distributives are quite productive. They might imply plurality of participants, but they do not specify it. In a few lexicalized nominals referring to persons, distributives now occur routinely with plurals, though not with duals. There are no collectives. There is little evidence of distinct verb roots or stems for different numbers of participants, or number suppletion for nouns.

There is no gender or number agreement within determiner phrases or noun phrases, unsurprising in light of the fact that articles and demonstratives do not distinguish number, and there is little evidence of either kind of phrase in any case. There is also no agreement within clauses, that is, no requirement that forms match

in gender or number. Many referring expressions were originally coined as verbs, with pronominal prefixes, such as *kaia'tákeras* 'goat', literally 'it-bodily-stinks' with singular zoic agent prefix *ka-*, but such terms have become lexicalized, and their forms remain unchanged no matter what the number of the referent in the clause.

2 Pronominal, nominal, and verbal number

2.1 Generalities

The Mohawk counterparts of unstressed pronouns in languages like English or Russian, which usually refer to given participants, are the pronominal prefixes on verbs. The Mohawk prefixes are fully referential, not simply markers of agreement with some other nominal in the clause. They differ from European unstressed pronouns in occurring in every clause, whether or not a coreferential lexical nominal or particle pronoun is also present. This is not a 'pro-drop' language, in which referents may be left unmentioned: speakers have clear intuitions that reference is specified overtly in every verb. Independent pronouns are used only for specific purposes, to indicate contrast of various types or to specify reference where there is no verb or noun to carry a prefix. The pronominal prefixes on verbs make more number distinctions than the independent pronouns.

Perhaps surprisingly, all basic morphological nouns are neuter in gender. They can occur, however, with possessive prefixes in place of the neuter gender prefix, and these distinguish a full range of gender and number for the possessors. Morphological verbs can also serve as lexical nominals, often without additional marking, referring to a core argument. Because they can contain the same sets of pronominal prefixes as other verbs, they can make similar gender and number distinctions. Kinship terms, formed from verbs describing relationships, exhibit full ranges of person, gender, and number for all members of the relationship. Finally, many particles, words with no internal morphological structure, have been lexicalized as referring expressions. These do not distinguish gender or number.

2.2 Pronominal number

The pronominal prefixes on verbs, and the possessive prefixes on nouns and other nominals, make more number distinctions than independent particle pronouns. Prefixes on basic morphological nouns, which are all grammatically neuter, make no number distinctions at all.

2.2.1 Pronominal prefixes on verbs

The basic forms of the pronominal prefixes that occur on verbs are in Table 1. Grammatical agent prefixes are in the column headed at the top by the symbol Ø. (The Ø indicates that no patient is mentioned in these forms.) The first person singular agent prefix (1SG.agt), for example, is *k-*, as in (1a). Grammatical patient prefixes are to the right of the symbol Ø in the left hand column. (The Ø indicates that no agent is mentioned in these forms.) The first person singular patient prefix (1SG.pat), for example, is *wak-*, as in (1b) Transitive prefixes can be found at the intersection of the agent row and the patient column. The basic form for a first person singular agent acting on a second person singular patient (1SG>2SG) is *kon-*. (All of the prefixes vary in shape with the preceding and following phonological context.)

(1) Sample pronominal prefixes[1]
    a. Agent
       Kká:wes.
       k-kawe-s
       1SG.agt-paddle-HAB
       '**I** paddle.'

    b. Patient
       Wakí:ta's
       wak-ita'w-s
       1SG.pat-sleep-HAB
       '**I** sleep.'

    c. Transitive
       Konhsá:tens.
       kon-hsaten-s
       1SG>2SG-carry.on.back-HAB
       '**I** carry **you**.'

The choice among agent, patient, and transitive paradigms is lexicalized with each verb stem. Usually the semantic basis behind the choice is still apparent, but in some cases it has become obscure. The transitive prefixes are fused forms: in some cases individual components can still be detected, but in others they cannot.

As can be seen in Table 1, singular, dual, and plural number are distinguished by the pronominal prefixes, but not in all forms. Among the agent prefixes, for

---

1 In the conventional orthography, most letters approximately represent their IPA values. The glide [j] is written <i>, and the nasalized vowels [ʌ] and [ų] are written <en> and <on> respectively. Glottal stop is written with an apostrophe <'>. Stress with high or rising tone is written with an acute accent <á>, and stress with sharply rising then falling tone is written with a grave accent <à>. Vowel length is indicated with a colon <a:>.

Tab. 1: Transitive Pronominal Prefixes. Basic Forms.

Agents＼Patients	1SG	1DU	1PL	2SG	2DU	2PL	Ø	N	FZ.SG	M.SG	M.DP	FI	F.DP
1SG				kon-	keni-		k-			-hri-	khe-		
1.EXC.DU							iakeni-			-hshakeni-			
1.EXC.PL				kwa-			iakwa-			-hshakwa-	iakhi-		
1.INC.DU							-eteni-			(eh)tshiteni-	iethi-		
1.INC.PL							-etewa-			(eh)tshitewa-			
2SG	tak-/-hsk-		takwa-/-hskwa-				-hs-			(eh)tsh-	-hshe-		
2DU	takeni-/-hskeni-						-eseni-			(eh)tshiseni-	ietshi-		
2PL							-esewa-			(eh)tshisewa-			
Ø	wak-	ionkeni-	ionkwa-	-esa-	-eseni-	-esewa-		ka/w-	io-	-hro-	-hroti-	iako-	ioti-
N							ka/w-						
FZ.SG													
M.SG	-hrak-	-hshonkeni-	-hshonkwa-	-hia-	(eh)tshiseni-	(eh)tshisewa-	-hra-			-hronwa-	-hshako-	-hshakoti-	
M.DU	ionk-	ionkhi-					-hni-				-hshakoti-/-hronwati-	iontat-	-hshakoti-/konwati-
M.PL							-hrati						
FI				iesa-	ietshi-		ie-		konwa-		-hronwati-		konwati-
F.DU							keni				-hronwati-/iakoti-	iakoti-	konwati-/iakoti-
F.PL							konti-						

1	FIRST PERSON ('I, we')	EX	EXCLUSIVE (excludes hearer)
2	SECOND PERSON ('you')	IN	INCLUSIVE (includes hearer)
M	MASCULINE ('he, they': males or mixed group)	DU	DUAL (two)
FI	FEMININE-INDEFINITE ('one, people, she')	PL	PLURAL (three or more)
N	NEUTER ('it, she': objects, animals, some women)	DP	DUO-PLURAL (two or more)
FZ	FEMININE-ZOIC (women, animals)		

example, three numbers are distinguished in all forms except feminine-indefinite and neuter. Among the patient and transitive prefixes, they are distinguished only in first and second persons. Third person patients, and those in transitive combinations, just distinguish singular and duo-plural. Some differences in number marking can be seen by tracing references to the couple in the passage below. The husband and wife are sometimes referred to with duals, and sometimes with duo-plurals. (Mixed groups, including couples, are classified as masculine.)

(2) Third person duals and duo-plurals: Mae Niioronhià:'a Montour, speaker p.c.
'This mother and father wanted to go out to a meeting.'
'They asked the grandmother to take care of the children while they were gone.'

> *Kí:ken **i**atathróna'*   *wà:**n**ehre'*
> kiken **i**-atat-hrona'   wa-**hn**-ehr-e-'
> this   **M.DU.AGT**-REFL-be.married   FAC-**M.DU.AGT**-want-EP-ST
> this   **they two** are married   **they two** wanted

> *a**hni**iá:ken'ne'*   *a**hi**atkennisa'ánha'.*
> aa-**hni**-iaken'n-e-'   aa-**hi**-atkennisa'-anha-'
> IRR-**M.DU.AGT**-go.out-EP-PFV   IRR-**M.DU.AGT**-meet-PURP-PFV
> **they two** would go out   **they two** would go to a meeting

> *Wa**hshakoti**ri'wanón:tonhse'*   *ne   akokstèn:ha*
> wa-**hshakoti**-ri'wanton-hs-e-'   ne   ako-ksten=ha
> FAC-**3DP>FI**-ask-BEN-EP-PFV   ART   FI.PAT-be.old=DIM
> **they** asked **her**   ART   she is old

> *aiontennonhna'*   *tsi   nikarì:wes*
> aa-ie-aten-nonhna-'   tsi   ni-ka-rihw-es
> IRR-FI.AGT-MID-watch-PFV   tsi   PRT-N.AGT-matter-be.long
> she would watch   as   so it is matter long

> *en**hoti**iakèn:'en.*
> en-**hoti**-iaken'-en
> FUT-**M.DP.PAT**-go.out-ST
> **they** will be out.

In the verbs 'they two are married', 'they two wanted', 'they two would go out', and 'they two would go to a meeting', the husband and wife are grammatical agents, so reference to them is dual. (In Mohawk, a distinction is made between inherent states, like 'be long', and resultant states, like 'be old'. Arguments of most inherent states are categorized grammatically as agents, and those of resultant states as patients.) In 'they will be out', the spouses are cast as duo-plural, because there is no dual distinction in third person patient prefixes. In 'they asked her', they are again cast as duo-plural, since there is no dual distinction in these third person transitive combinations. Comparative evidence indicates that while the singular/dual/plural

distinction for first and second persons goes back to Proto-Iroquoian, duals for third persons have emerged more recently in the Northern Iroquoian languages.

In general, number is distinguished only for animates. Neuter prefixes have nearly the same forms as feminine-zoics, but, importantly, they never distinguish number. (Neuters are also not mentioned overtly in the pronominal prefix if another argument is present. As can be seen in Table 1, for example, the prefix *k-* for 'I>it' is the same as that simply for 'I'. The prefix *wak-* is both 'it>me' and simply the patient 'I'.) In (3), the same pronominal prefixes are used to refer to one or more inanimate objects, whether they are grammatical agents or patients.

(3) Neuters
   ***ka**-rihwén:ta's*   '**it** wears out, **they** wear out' N AGT
   ***io**-ra'nentá:kon*   '**it** is stuck, **they** are stuck'   N PAT

Number is usually distinguished for feminine-zoics, however, that is, in prefixes used to refer to animals and some female persons. The differences can be seen in (4).

(4) Feminine-zoics
   ***ká**-tie's*   '**it** (a bird) flies around'   F SG.AGT
   ***kení**-tie's*   '**they** (two birds) fly around'   F.DU.AGT
   ***kontí**-tie's* '**they** (three or more birds) fly around' F.PL.AGT

   ***io**-iehwá:ton*   '**it** (a bird) has woken up'   F.SG.PAT
   ***ioti**-iehwá:ton* '**they** (two or more) have woken up' F.DP.PAT

Intriguingly, however, certain verbs do distinguish number for referents that might ordinarily be considered inanimate. In many cases, a rationale can be imagined. The passage in (5) comes from a description of planting corn. In the verbs 'sprout', 'pierce the soil', 'care for', and 'grow', the corn plants are treated as animate, with pronominal prefixes that distinguish number. The free English translation was proposed by the speaker.

(5) Plural corn plants: Josephine Kaieríthon Horne, speaker p.c.
   Tsahià:kshera   *sok nà:'a eniotikè:tohte',*
   ts-iahia'k-hser-a-t   sok   nà:'a en-**ioti**-ke'toht-e-'
   REP-cross-NMLZ-LK-be.one then guess FUT-**FZ.DP.PAT**-appear-PFV
   one week   then guess **they** will emerge
   'In about a week's time we would see the tiny green blades sprout,'

   kén'   nikanerahté:son's   *tenkonta'kenhrótka'we'.*
   ken'   ni-ka-neraht-es-on's   t-en-**kon**-at-a'kenhr-otka'w-e-'
   small PRT-N-leave-be.long-DISTR DV-FUT-**FZ.PL.AGT**-MID-soil-pierce-EP-PFV
   small so it are sized variously **they** will soil pierce
   'come through the ground.'

*É:so' ten**konwatí**hsnie'ne'*
é:so'   t-en-**konwati**-shnie'n-e-'
much   DV-FUT-**FI>3PL**-care.for-EP-PFV
much   **one** will care for **them**
'After much care and nurturing'

*kháre' ó:nen en**kont**ehiá:ron'.*
kháre' ó:nen en-**kon**-ate-hiaron-'
until   now   FUT-**FZ.PL.AGT**-MID-grow-PFV
until   now   **they** will grow
'they would grow tall and stately.'

On another occasion, another speaker commented on the crops seen on a drive through the countryside.

(6) Plural corn: Watshenní:ne' Sawyer, speaker p.c.
    *Ó:nenhste'   ken'k   ni**konti**hnenié:son's,*
    o-nenhst-e-'   ken'=k   ni-**konti**-hneni-es-on's
    N-corn-EP-NS small=only   PRT-**FZ.PL.AGT**-length-be.long-DISTR
    corn   small   so **they** are each long
    'The corn (plants) are very short,'

    *tsi ní:io   nakwé: thie-**ioti**:-ten.*
    tsi ni-io-ht   ne=akwek-on th-ie-**ioti**-iten
    as PRT-N.PAT-be.so   ART=all-ST   CONTR-TRL-**FZ.DP.PAT**-be.poor
    as so it is   the all   **they** are just poor there
    'They all seem to be doing poorly.'

Speakers agree that this is not simply a stylistic choice, but the idiomatic way to present these situations. As one remarked, 'No one would ever say *Ken'k niión:son's* 'it is short' referring to corn.'

    Plural forms also appear with some other mentions of plants. In (7), the speaker referred to her flowers in the plural in the verb 'I would water them', but not in the verb 'I have planted (them)'. (As noted above, neuters are not identified overtly in pronominal prefixes if another argument is present.)

(7) Plural flowers: Dorothy Karihwénhawe' Lazore, speaker p.c.
    *Wà:kehre'   en**khe**hnekanontèn:ra'*
    wa'-k-ehr-E'   en-**khe**-hnek-a-nonten-hra-'
    FAC-1SG.agt-think-PFV FUT-**1SG/3pl**-water-JOINER-feed-PURPOSIVE-PUNCTUAL
    I thought   I would go water feed **them**
    'I thought I would water'

kí:ken ohén:ton' tewaktsi'tsiaiénthon.
kí:ken ohén:ton' te-wak-tsi'tsi-a-ient-hw-on
this in.front CSL-1SG.pat-flower-LK-lie-CAUS-ST
this in front there I have flower planted
'the flowers I had planted in front of the house.'

It should be noted that incorporated nouns are not themselves arguments: they simply qualify the verb, narrowing its scope semantically. The resulting complex verb can be intransitive or transitive; some such verbs are simply lexicalized as intransitives, some as transitives, and some as both.

In (8), plurality was distinguished in the verb describing ripe potatoes.

(8) Plural potatoes: Kanerahtenhá:wi Hilda Nicholas, speaker p.c.
Ion*atonníson'.*
ion-at-onni-is'-on
FZ.DP.PAT-MID-make-CMPL-ST
they have finished making themselves
'**They** (potatoes) are ripe.'

But it was not distinguished in the verb describing ripe berries.

(9) Berries: Kanerahtenhá:wi Hilda Nicholas, speaker p.c.
Io*hiá:ri.*
io-ahi-ari
N.PAT-berry-ripe.ST
'**They** are ripe (in the way that berries are ripe).'

Two different verbs were used here for being ripe. The one used with potatoes is literally 'having finished making oneself', i.e. growing, while that used with berries is simply 'be ripe, cooked'.

Some additional verbs pertaining to plants that have occurred with number specification in their pronominal prefixes are in (10). They contain the feminine-zoic duo-plural patient prefix *ioti-*.

(10) Verbs pertaining to plants
      ioti*-ká:ion*   '**they** are growing slowly' (corn)   DP
      ioti*-kenhé:ion* '**they** are slow growing' (corn, berries, potatoes, grass) DP
      ioti*-hniò:'on* '**they** just grew'   (volunteer trees in the woods) DP

Speakers agree that number specification is obligatory with these verbs.

Plurality is also sometimes specified in pronominal prefixes on verbs for objects that move, like trains and airplanes. The plurality is not distinguished on the associated nominals.

(11) Plural trains: Joe Tiorhakwén:te' Dove, speaker p.c.

Thí:kén tio'kéha' se's non:,
thiken teio'kéha' se's nón:we
that train formerly place
'Those trains, you know,'

*te**kontí**ta's.*
te-**konti**-t-a'-s
DV-**FZ.PL.AGT**-stand-INCH-HAB
'**they** used to stop there.'

It might be tempting to conclude that the terms for 'corn plants', 'flowers', 'potatoes', and 'trains' are covertly specified for feminine-zoic gender and number, and that the pronominal prefixes on verbs simply agree with them. This is not a matter of agreement, however, nor is it even a property of the noun. Example (12) is from an account of how the king of France sent a bell to the Kahnawà:ke community in Quebec for their church. He put it on a ship, but it never arrived at its destination. The ship was captured by the English and taken to Massachusetts. At this point, the ship was referred to as zoic in the pronominal prefix on the verb 'they caught it', but as neuter on the verbs 'they took it' and 'it was captured'.

(12) Zoic and neuter ship: Josephine Kaieríthon Horne, speaker p.c.

*Wa'**konwa**ié:na' ne **ka**honweia'kó:wa,*
wa'=**konwa**-iena-' ne ka-honwei-a'=kowa
FAC-3PL>**FZ.SG**-capture-PFV ART N-boat-NS=AUG
they caught **it** the ship
'They captured the ship (**FEMININE-ZOIC**),'

iahatíhawe' wastonhronòn:ke.
i-a-hati-haw-e-' waston=hronon'=ke
TRL-FAC-M.PL.AGT-take-EP-PFV Boston=resident=place
they took **i**t Boston place
'and took it (**NEUTER**) to the United States.'

*Deerfield tkonwá:iats, tsi nón: ti**o**ia'totarhè:'on.*
NAME t-konwa-iat-s tsi nonhwe' t-**io**-ia't-otarhe'-on
CSL-3PL>FZ.SG-call-HAB at place CSL-**N.PAT**-body-hook-ST
Deerfield they call it (FEMININE-ZOIC) at place there **it** (**NEUTER**) was caught
'It (**NEUTER**) was captured at a place called Deerfield.'

At Deerfield the cargo was sold, and the town bought the bell for their church. When the Mohawks heard of this, they set out to reclaim their bell. Once there, they took the bell down from the church steeple and fastened it to a beam to carry back home. At this point the bell was referred to as neuter in the verbs 'they tied it' and

'they carried it'. (Recall that neuters are not mentioned overtly in pronominal prefixes if another argument is present.)

(13) Neuter bell: Josephine Kaieríthon Horne, speaker p.c.
      *Karontà:ke   wa**hati**hwánerenke'*.
      ka-ront-a'ke wa-**hati**-hwanerenk-e-'
      N-log-place   FAC-**M.PL.AGT**-tie-EP-PFV
      log place   they tied it
      'They fastened it (**NEUTER**) to a beam.'

      *Wa**hon**nentshà:ren'*   ...
      wa-**hon**-a-nentsh-a-hren-'
      FAC-**M.PL.AGT**-MID-arm-LK-set.on-PFV
      they set it on their arms
      'They carried it (**NEUTER**) on their shoulders ... '

It was winter, however, and travel was difficult, so they decided to bury the bell and return for it in the spring. At this point the bell was categorized as feminine-zoic in the pronominal prefix on the verb 'they would bury it'.

(14) Feminine-zoic bell: Josephine Kaieríthon Horne, speaker p.c.
      *En**konwa**ia'táta'*   *ne   iehwista'ékstha'*.
      en-**konwa**-ia't-a-t'a-'   ne   ie-hwist-a-'ek-st-ha'
      FUT-3PL**>FZ.SG**-body-LK-insert-PFV ART   FI.AGT-metal-LK-strike-INS.APPL-HAB
      they will bodily insert **it**   ART   one metal strikes with it
      '[They decided] they would bury the bell (**FEMININE-ZOIC**).'

It is the verb that controls the classification of the argument as neuter (unmentioned in transitives, and with no number distinctions) or feminine-zoic (with number). Relatively few verb stems are used with both inanimate and animate referents: most occur with only one or the other. A few press toward conception of participants as animate, like those seen here with meanings 'sprout', 'grow', 'feed', 'stop', 'catch', 'bury', and 'call by name'. The primary principle at work here is the classification of events and states as pertaining to animate or inanimate participants, not to the direct classification of entities, but the classification can have implications for number marking.

2.2.2 Independent pronouns

There are no independent pronouns in Mohawk comparable to the unstressed pronouns of languages like English or Russian: their counterparts are the pronominal prefixes on verbs. The independent pronouns that do exist have special uses, signal-

ing contrast or emphasis as in (15) below, or occurring when there is no verb in the sentence for a pronominal prefix to attach to as in (16). These independent pronouns do not replace the pronominal prefixes on associated verbs.

(15) Independent pronoun: Margaret Edwards, speaker p.c.
Akwé:kon átste'   *wa'akwákien*   *wa'kiakwatskà:hon'*
akwekon atste'   wa'-iakwa-at-ien'   wa'-t-iakwa-at-ska'nhon-'
all   outside FAC-1PL.EX.AGT-MID-set-PFV FAC-DV-1PL.EX.AGT-MID-dine-PFV
'We all sat outside and ate.'

Sok **nì:'i** *wa'akwatkahri'tsherón:ni,*
sok **ne=i'i** wa'-iakwa-at-kahri-'tsher-onni-'
then **ART=1** FAC-1PL.EX.AGT-MID-amuse-NMLZ-make-PFV
'Then **us**, we'd play,'

iakwaksa'okòn:'a.
iakwa-ksa'=okon'=a
1PL.EX.AGT-be.child=DISTR=DIM
'us kids.'

(16) Independent pronoun: Watshenní:ne' Sawyer, speaker p.c.
*Ónhka' n**í:se**'?*
onhka' ne=**ise'**
who   ART=**2**
'Who are **you**?'

The independent pronouns do not distinguish grammatical role: they are used for agents or patients, and inalienable or alienable possessors. The first and second person forms, which are older than the third person forms, do not distinguish number or clusivity: first person *ì:'i* can be translated 'I myself', 'me', 'my', 'we', 'us', or 'our'. The third person forms distinguish both gender and singular/duo-plural number when referring to persons.

(17) Independent pronouns
ì:'i   1
í:se' 2

aónha'   3 FEMININE-ZOIC SINGULAR
akaónha' 3 FEMININE-INDEFINITE SINGULAR
onónha'   3 FEMININE DUO-PLURAL

raónha'   3 MASCULINE SINGULAR
ronónha' 3 MASCULINE DUO-PLURAL

A masculine duo-plural pronoun in speech is in (18).

(18) Independent pronoun: Josephine Kaieríthon Horne, speaker p.c.
'They envied the Indians who were wearing snowshoes
and seemed to be floating on top of the snow.'

Wahonská:neke' **ronónha** *ò:ni'*
wa-hon-askanek-e-' **rononha** ohni'
FAC-M.PL.AGT-wish-EP-PFV **M.DP** also
'They wished that they **themselves**'

ahotiién:take' *ne kahwèn:kare'.*
aa-hoti-ient-ak-e-' ne ka-hwen'kar-e-'
IRR-M.DP.PAT-have-CONT-EP-PFV ART N-board-EP-NS
'had snowshoes too.'

2.2.3 Demonstrative, indefinite, and interrogative pronouns

The two demonstrative pronouns *kí:ken* 'this one, these', and *thí:ken* 'that one, those' serve as referring expressions on their own, though they may occur in apposition with lexical nominals. They do not distinguish number. In (19), the proximal *kí:ken* is referring to a single man. In (20), the same demonstrative is referring to multiple ants. It is clear that just one man is under discussion in (19) from the masculine singular pronominal prefix on the verb 'he arrived', and that multiple ants are under discussion in (20) from the masculine plural pronominal prefix on the verb 'they set up a ladder'.

(19) Demonstrative: Annette Kaia'titáhkhe' Jacobs, speaker p.c.
Iahà:rawe' **kí:ken,**
iah-a-hra-w-e-' **kiken**
TRL-FAC-M.SG.AGT-arrive-EP-PFV **this.one**
'He arrived, **this guy**,'

a'è:ren ní:wa' *iokà:ronte'.*
a'eren ni-w-a' io-kahront-e-'
over PRT-N.AGT-be.a.size N.PAT-be.a.hole-EP-ST
'at a big canyon.'

(20) Demonstrative: Dorothy Karihwénhawe' Lazore, speaker p.c.
Ronatenekó:tote' **kí:ken** *tsiki'nhontstókhi'.*
ron-ate-nekot-ot-e-' **kiken** tsiki'nhontstokhi'
M.PL.AGT-MID-ladder-stand-EP-PFV **these** ant
'**These** ants had set up a ladder.'

Similarly, indefinite pronouns like 'someone' and interrogative pronouns like 'who' make no gender or number distinctions, even when referring to animates. The ques-

tion in (16) above, *Ónhka' ní:se'* 'Who are you?' could be directed at one or more persons. That in (21) below could refer to one or more.

(21) Interrogative and demonstrative pronouns
      *Ónhka' **thí:ken**?*
      who   **that**
      'Who is that?', 'Who are those people?'

2.2.4 Reflexives

There are no reflexive pronouns: intransitive reflexive verb bases are formed with a verb prefix *-atat-*. (Forms vary with phonological context.)

(22) Reflexive construction: Mae Niioronhià:'a Montour, speaker p.c.
      *Wa'**ontat**kwatá:ko'.*
      wa'-ie-**atat**-kwatako-'
      FAC-FI.AGT-**REFL**-prepare-PFV
      'She prepared **herself**' = 'She got ready.'

(23) Reflexive construction: Carolee Konwatién:se' Jacobs, speaker p.c.
      *Enk**atat**iénhahse'.*
      en-k-**atat**-ien-hahs-e-'
      FUT-1SG.agt-**REFL**-have-BEN-EP-PFV
      'I will save it for **myself**.'

The reflexive prefixes show no number distinctions in themselves. Number is of course indicated in pronominal prefixes.

2.3 Nominal number

All kinds of words can be used as referring expressions or lexical nominals, not just basic morphological nouns, but also morphological verbs and particles. The number distinctions made in each depend to some extent on their origins.

2.3.1 Basic nouns

Basic morphological nouns have relatively simple structures. They consist of a prefix indicating the gender of the noun or of the person, gender, and number of a possessor, then a noun stem, and then a noun suffix. (The noun suffix adds no meaning apart from marking the word as a noun.) It appears that all basic nouns

are neuter. They thus do not distinguish number. The most common neuter prefixes on nouns are *o-* and *ka-*, with phonologically-conditioned zero allomorphs before the oral vowels *a* and *e,* and the nasal vowels *en* and *on.*

(24)   Basic nouns
  a. *okà:ra'*
     o-kahr-a'
     N-eye-NS
     'eye(s)'

  b. *o'wháhsa'*
     o-'whahs-a'
     N-skirt-NS
     'skirt(s)'

  c. *kátshe'*
     ka-tsh-e'
     N-jug-NS
     'jug(s), bottle(s), can(s)'

  d. *áhsire'*
     ahsir-e-'
     blanket-EP-NS
     'blanket'

  e. *onhwentsa'*
     onhwentsi-a'
     earth-NS
     'earth'

The noun prefixes *o-* and *ka-* are apparently descended from neuter pronominal prefixes on verbs, but on nouns, there is now no semantic difference between the two (Mithun to appear).

The noun stem may be derived, formed with a nominalizing suffix.

(25)   Derived nouns
  a. *owistóhsera'*
     o-wisto-hser-a'
     N-be.cold-NMLZ-NS
     'butter'

  b. *kahiatonhsera'*
     ka-hiaton-hser-a'
     N-write-NMLZ-NS
     'paper(s), book(s)'

c. *onentshawì:ta'*
o-nentsh-awi-'t-a'
N-arm-encircle-NMLZ-NS
'wrist(s)'

Nouns may begin with a possessive prefix in place of the neuter prefix. Alienable and inalienable possession are distinguished. Most body parts, particularly those with an outside surface, are inalienably possessed. These usually occur with a derivational place ending, which derives terms for locations. Most kinship terms are not referred to as possessions, but as relationships, described in section 2.3.3.

(26) Alienably possessed nouns
a. ***ake****'wháhsa'*
ake-'whahs-a'
1SG.AL.POSS-skirt-NS
'**my** skirt(s)'

b. ***ak****hiatónhsera'*
ak-hiaton-hser-a'
1SG.AL.POSS-write-NMLZ-NS
'**my** paper(s), book(s)'

(27) Inalienably possessed nouns
a. ***k****kahrà:ke*
k-kahr-a'ke
1SG.INAL.POSS-eye-place
'**my** eye(s)'

b. ***ke****nentshawi'tà:ke*
ke-nentsh-awi-'t-a'ke
1SG.INAL.POSS-arm-encircle-NMLZ-place
'(on) **my** wrist(s)'

The nouns themselves, as neuters, do not distinguish number. The same nouns are used for 'skirt' and 'skirts', and for 'eye' and 'eyes'. The possessive prefixes, however, distinguish the number of animate possessors. Many of the possessive prefixes are quite similar in form to pronominal prefixes on verbs. The first and second person *alienable* possessive prefixes resemble the *patient* prefixes on verbs, and the first and second person *inalienable* prefixes resemble the *agent* prefixes on verbs. (Where the verbal prefixes begin with a glide, their counterparts on nouns lack the glide.) The possessive prefixes make the same number distinctions as their counterparts on verbs: singular/dual/plural for first and second persons; singular/duo-plural for third person alienable possessors; and singular/dual/plural for inalienable third person possessors.

(28) Alienably possessed nouns
  a. *akhiatónhsera'*
     ak-hiaton-hser-a'
     1SG.AL.POSS-write-NMLZ-NS
     'my paper(s), book(s)'

  b. *onkenihiatónhsera'*
     onkeni-hiaton-hser-a'
     1DU.AL.POSS-write-NMLZ-NS
     'we two, our paper(s), book(s)'

  c. *onkwahiatónhsera'*
     onkwa-hiaton-hser-a'
     1PL.AL.POSS-write-NMLZ-NS
     'we (three or more), our paper(s), book(s)'

  d. *raohiatónhsera'*
     rao-hiaton-hser-a'
     M.SG.AL.POSS-write-NMLZ-NS
     'his paper(s), book(s)'

  e. *raonahiatónhsera'*
     raona-hiaton-hser-a'
     M.DP.AL.POSS-write-NMLZ-NS
     'their (two or more) paper(s), book(s)'

(29) Inalienably possessed nouns
  a. *kkahrà:ke*
     k-kahr-a'ke
     1SG.INAL.POSS-eye-place
     'my eye(s)'

  b. *tenikahrà:ke*
     teni-kahr-a'ke
     1INCL.DU.INAL.POSS-eye-place
     'the two of us, our eyes'

  c. *tewakahrà:ke*
     tewa-kahr-a'ke
     1IN.PL.INAL.POSS-eye-place
     'our (three or more) eyes'

  d. *rakahrà:ke*
     ra-kahr-a'ke
     M.SG.INAL.POSS-eye-place
     'his eye(s)'

  e. *nikahrà:ke*
   ni-kahr-a'ke
   M.DU.INAL.POSS-eye-place
   'those two, their eyes'

  f. *ratikahrà:ke*
   rati-kahr-a'ke
   M.PL.INAL.POSS-eye-place
   'their (three or more) eyes'

Though basic nouns contain no indication of number, distributive enclitics can be added to them and other nominals that can imply number. A distributive can be seen in (30) below. This enclitic distributes entities over types. Thus *otsikhe'ta'shòn:'a* cannot refer simply to multiple candy canes; the candies must be of different kinds. The enclitic itself does not specify plurality, but it can imply it.

(30) Nominal distributive
  a. *otsikhè:ta'*
   o-tsikhe't-a'
   N-sugar-NS
   'sugar, candy'

  b. *otsikhe'ta'shon:'a*
   o-tsikhe't-a'=shon'a
   N-sugar-NS=DISTR
   'various and assorted candies' (lollypops, candy canes, gumdrops, etc.)

There are no collective or associative markers.

  Enumeration of entities is accomplished with verbs. To specify that there is just one inanimate object, the noun stem for that object can be incorporated into a verb *-at* meaning 'be one in number' with prefixed repetitive marker *s-*. (In most dialects the final *t* of the verb *-at* is no longer pronounced.) To specify that there are two objects, the noun stem is incorporated into a verb *-ke* 'be multiple' with prefixed duplicative marker *te-*. To specify that there are more, the noun stem is incorporated into the same verb with prefixed partitive marker *ni-*, the whole word then preceded by the numeral.

(31) Enumeration
  a. *ska'wháhsa*
   s-ka-'whahs-at
   REP-N.AGT-skirt-be.one
   'one skirt'

b. *teka'wháhsake*
   te-ka-'whahs-a-ke
   DV-N.AGT-skirt-LK-be.multiple
   'two skirts'

c. *áhsen nika'whahsake*
   ahsen ni-ka-'whahs-a-ke
   three PRT-N.AGT-skirt-LK-be.multiple
   'three skirts'

Enumeration is a verbal construction though, like other verbs, these verbs can be used to refer.

The verb root in enumeration constructions need not be *-at* 'be one' or *-ke* 'be multiple'. Enumerative constructions can be formed from other verbs with the duplicative prefix on the verb or a numeral and partitive prefix on the verb.

(32) Enumeration: Sha'tekenhátie' Marian Phillips, speaker
    Ià:ia'k   *na'kahwistà:'eke'*
    iahia'k   **n**-a'-ka-hwist-a-'ek-e'
    cross.over PRT-FAC-N.AGT-bell-LK-strike-PFV
    six   **so** many it bell struck
    'For **six hours**'

    tiio'kehà:ke   *ionkwahonwì:sere'.*
    teio'keh-a'=ke ionkwa-honw-i'ser-e'
    train-NS=place 1DP.PAT-conveyance-drag-HAB
    train place   it dragged us
    'we rode on the train.'

If an object is already under discussion, it need not be mentioned with a full noun. In its place the verb *ní:kon* 'so it amounts to' can be used.

(33) Enumeration: John Maracle, speaker
    Wísk ní:kon   *enhsohará:ko'.*
    wisk ni-ka-on   en-hs-ohar-ako-'
    five PRT-N.AGT-amount FUT-2SG.agt-attach-REV-PFV
    five   so it amounts   you will detach
    'You'd cut up **five of them**.'

When the entities enumerated are animate, the constructions are slightly different. For one, the noun root *-ia't-* 'body' is incorporated into the verb *-at* 'be one'.

(34) One person or animal
     shaià:ta
     s-ha-ia't-at
     REP-M.SG.AGT-body-be.one
     'one male person'

For two, the verb root *-iahs-* 'be a pair' is used with the duplicative verb prefix *te-*.

(35) Two persons or animals
     tehniiáhse
     te-hni-iahse
     DV-M.DU.AGT-be.a.couple
     'two men or boys, a couple'

For three or more, the verb root *-i* is used with the partitive prefix *ni-* and a numeral.

(36) Three or more
     áhsen nihá:ti
     ahsen ni-hati-i
     three   PRT-M.PL.AGT-number
     'three men, boys, or people'

Of course all of these verbs contain pronominal prefixes specifying number.

There are a few nouns referring to groups, but they behave like regular basic neuter nouns. As neuters, they mark no overt number distinctions themselves.

(37) Groups
     a. *kentióhkwa'*
        ka-itiohkw-a'
        N-group-NS
        'group(s)'

     b. *kahwà:tsire'*
        ka-hwa'tsir-e-'
        N-family-EP-NS
        'family, families'

There is no distinction between mass and count nouns. Mass nouns have the same forms as others, like *ó:nenhste'* 'corn', with neuter prefix *o-*.

(38) *ó:nenhste'*
     o-nenhst-e-'
     N-corn-EP-NS
     'corn'

(These are of course not enumerated with constructions like those in (34)–(36). Since number is unmarked for neuters, there are no *dualia* or *pluralia tantum* nouns. The term for 'pants', for example, is a simple basic noun *atháhsteren* with zero neuter prefix before the initial *a*.

Since there are no number distinctions for basic nouns, which are all neuter, there is no number-based suppletion.

2.3.2 Verbal nominals

As noted earlier, essentially all basic nouns are neuter. This seems at first surprising. Surely there are ways of designating people. There are indeed, but apart from a noun for 'baby', *owirà:'a* (which is grammatically neuter), terms for people do not have the full characteristics of basic morphological nouns. Morphological verbs are often used as referring expressions syntactically, and they can become lexicalized as nominals, often with no further marking. But like other verbs, they can contain pronominal prefixes which distinguish number for animate referents.

(39) Verbal nominals
   a. ***ra'swà:tha'***
      ra-a'swa't-ha'
      M.SG.AGT-extinguish-HAB
      '**he** extinguishes' = 'fireman'

   b. ***ron'swà:tha'***
      ron-a'swa't-ha'
      M.PL.AGT-extinguish-HAB
      '**they** extinguish' = 'firefighters'

Only morphological noun stems can be incorporated into verbs. Any other terms must first be overtly nominalized with a nominalizing suffix before being incorporated. (The nominalizer may or may not be present when the term is used as a separate word.) Terms for persons other than babies are either not incorporated at all, or are incorporated only with a nominalizer.

(40) Children
   a. *raksà:'a*
      ra-ksa'=a
      M.SG.AGT-be.a.child=DIM
      'boy'

   b. *raksa**'tí**:io*
      ra-ksa-**'t**-iio
      M.SG.AGT-be.a.child=**NMLZ**-be.good
      'he is a good boy' = 'good boy'

(41) Persons
   a. *rón:kwe*
      r-onkwe
      M.SG.AGT-be.a.person
      'man'

   b. *ronkwe'tí:io*
      r-onkwe-**'t**-iio
      M.SG.AGT-be.a.person-**NMLZ**-be.good
      'he is a good person' = 'good man'

Their prefixes distinguish number.

(42) Multiple persons
   a. *te**hni**ksà:'a*
      te-**hni**-ksa'=a
      DV-**M.DU.AGT**-be.a.child=DIM
      'two boys'

   b. ***rati**ksa'okon:'a*
      rati-ksa'=okon'=a
      M.PL.AGT-be.a.child=DISTR=DIM
      'boys, children (three or more)'

The dual form also occurs with the duplicative prefix *te*-, a more recent addition reinforcing the duality. The plural form contains a distributive. Not surprisingly, these particular formations are highly frequent and have become strongly lexicalized. Though in principle distributives are not obligatory, occurring only to distribute entities over types, the term for plural boys or children (though not the dual) always occurs with this distributive ending.

   Large numbers of terms for persons and animals have been formed from morphological verbs. Many have become lexicalized labels, recognized primarily as referring expressions. Among them are terms for animals. The term for 'cow', for example, is descriptive of its protruding jowls. It should be noted that the duplicative prefix pertains to the jowls, as part of the description, and not to the referent of the whole, the cow(s).

(43) Verbal nominal
   *te**io**nhónhskwaron*
   te-**io**-nhonhskwar-ont
   DV-**FZ.PAT**-jowl-be.attached
   it jowl protrudes doubly
   'cow'

The very same term is used for multiple cows, with no change in the prefix. As elsewhere, number is still distinguished in the pronominal prefixes on verbs that refer to cows, but the term itself for cows is inert.

(44) Cow mismatch: Rita Konwatsi'tsaién:ni Phillips, speaker p.c.

Ísi'	*nónhskwati*		*she's*	*niió:re'*
isi'	ne=w-ahskw-a-ti		she's	ni-io-r-e'
yonder	ART=N.AGT-bridge-LK-be.beyond		formerly	PRT-N.PAT-extend-ST
yonder	bridge beyond		formerly	so it is far

'Over there beyond the bridge,'

*nie***konn***éhtha'*		*ne*	*te***io***nhónhskwaron.*
n-ie-**konn**-e-ht-ha'		ne	te-**io**-nhonhskwar-ont
PRT-TRL-**FZ.PL.AGT**-go-DIR-HAB		ART	DV-**FZ.PAT**-jowl-be.attached
there **they** (**FZ**) would go to		ART	it (**NEUTER**) double jowl protrudes

'the cows (**no number**) used to pasture (**PLURAL**).'

A similar apparent number mismatch between nouns and verbs can be seen in (45). The prefix on the term for 'big bugs', a verbal nominal based on the verb root 'be big', contains no indication of number, but the pronominal agent prefix on the verb 'they will fly' is plural.

(45) Bug mismatch: Dorothy Karihwénhawe Lazore, speaker p.c.

Kaniataratátie'	*en***kont***ohétstake'*	*ka***tsi'noniowá:nen's*
ka-niatar-atatie'	en-**kon**-tohetst-ak-e-'	**ka**-tsi'nonw-owan-en-'s
N-river-along	FUT-**FZ.PL.AGT**-pass-CONT-EP-PFV	**N**-bug-be.big-ST-DISTR
along the river	**they** will pass by	**it** is each bug large

'Insects (**no number**) will pass by (**PLURAL**) along the river'

ok	*ò:ni'*	*ne*	*karonhia'kéhshon'*	*en***kont***ítie'*
ok	ohni'	ne	ka-ronhi-a'=ke=hshon'	en-**konti**-tie'
and	also	ART	N-sky-NS=place=DISTR	FUT-**FZ.PL.AGT**-fly-PFV
and	also	ART	through the skies	**they** will fly

'and they will also fly (**PLURAL**) through the skies'

se'ken	*ne*	*ka***tsi'noniowá:nen's.*
se'ken	ne	**ka**-tsi'nonw-owan-en-'s
also	ART	**N**-bug-be.large-ST-DISTR

'as well, those insects (**no number**).'

It is not always the singular form that is lexicalized as the basic term. A term for 'sheep' was also coined as a verb. It contains a plural prefix, which remains in place no matter how many sheep are referred to.

(46) Sheep
teiotina'karontòn:'a
te-**ioti**-na'kar-ont=on'=a
DV-**FZ.DP.PAT**-horn-be.attached=DISTR=DIM
'sheep' (any number)

When just one is specified explicitly, the prefix remains unchanged.

(47) One sheep
skaià:ta   *teiotina'karontòn:'a*
s-ka-ia't-at   te-**ioti**-na'kar-ont-on'-a
REP-N-body-be.one DV-**FZ.DP.PAT**-horn-be.attached-DISTR=DIM
one body   **they** are horn attached
'one sheep'

One could imagine that since sheep most often occur in flocks, it was the plural form that was most frequent and that became established. In the eastern Mohawk dialects, there is another term for 'sheep': *timotón*. This is clearly borrowed from the French *des moutons*, which is itself also plural. The Mohawk *timotón* is used for any number of sheep, one or more.

Terms for abstract concepts are sometimes formed with overt nominalization, but such formations occur most often in translations from French or English.

(48) Nominalization: Josephine Kaieríthon Horne, speaker p.c.
attokháhtshera'
w-at-tokha'-htsher-a'
N-MID-notice-NMLZ-NS
'wisdom'

More often regular finite verbs and clauses are used. Actions are not normally referred to with nominalizations. Full clauses are used instead.

(49) Clause: Josephine Kaieríthon Horne, speaker p.c.
Ákte' nonsakaié:ra'te'   *tsi* **ronaterí:io**.
akte' n-onsa-ka-ier-a't-e-'   tsi **ron-ate-riio**
other PRT-FAC.REP-N.AGT-do-CAUS-EP-PFV as **M.DP.PAT-MID-fight**
other it turned back   as **they are fighting**
'**The battle** turned [and they won].'

(50) Clause: Billy Kaientarónkwen Two Rivers, speaker p.c.
*Iáh ki' **tesewatia'tarohròn:ne'**.*
iah ki' **te-sewa-at-ia't-a-rohr-on-hne'**
not in.fact **NEG-2PL-MID-body-LK-gather-ST-PAST**
not in fact **did you gather yourselves together**
'You didn't have **a gathering**.'

2.3.3 Kinship terms

In general, kinsmen are not referred to in Mohawk as possessions, but rather in terms of relationships. They were coined as verbs. Thus one does not say 'my grandfather', but rather 'he is grandparent to me'. The senior relative is identified like the grammatical agent in the pronominal prefix, and the junior relative like the grammatical patient. They show the same number distinctions as verbs for all parties in the relationship.

(51) Kinship
ronwahsótha
ronwa-hsot=DIM
M.DP>M.SG-be.grandparent.to=DIM
'**they** are grandparents to **him**' = 'his grandparents'

(52) Kinship
konwa'kèn:'a
konwa-'ken'=a
3DP>FZ.SG-have.as.younger.sibling=DIM
'**they** have **her** as younger sibling' = 'their younger sister'

Reciprocal relationships are based on verbs formed with reflexive/reciprocal prefixes that follow the intransitive pronominal prefixes. The forms can refer to all members of the relationship or just some.

(53) Reciprocal relationships
iatate'kèn:'a
hi-**atate**-'ken'=a
M.DU.AGT-**REFL**-have.as.sibling=DIM
'they two have **each other** as siblings'
= 'those two siblings', 'his brother', 'her brother', 'his sister'

(Most kinship terms contain diminutive endings in referring forms, presumably originally as a marker of affection.) The pronominal prefixes on kinship terms are essentially the same as those on verbs, but some, like nouns, lack initial glides.

2.3.4 Particle nouns

Other words that can function as nominals are morphological particles, words with no internal structure. The term for 'chicken', for example, was apparently borrowed during the 17th century from Dutch, when Mohawks heard Dutch settlers calling

their chickens. These nominals do not distinguish number, no matter what their gender. In (54) the chickens are the grammatical agents of the clause. There is no number marking on *kítkit* 'chicken', but it is clear that multiple chickens were involved from the pronominal prefix on the verb 'they are speaking'.

(54) Particle noun as agent: Ima Johnson, speaker

Nen	*kati'*	*'eh*	*non's*	*tei**otí**hthare'*		*thí:*	***kítkit.***
onen	kati'	'eh	nonwe's	te-**ioti**-hthar-e-'		thiken	**kitkit**
now	in.fact	there	place=DISTR	DV-**FZ.DP.PAT**-speak-EP-ST		that	**chicken**
then	in fact	there	here and there	**they** are talking		that	**chicken**

'So then, at that place, those **chickens** were going around clucking.'

In (55), the chickens are the grammatical alienable possessors of the chicken feed. Again the word *kítkit* 'chicken' is unmarked for number, but the possessive prefix on 'their food' is plural.

(55) Particle noun as possessor: Ima Johnson, speaker

Enhshní:non'	*ne*	*kítkit*	***aotí**khwa'.*
en-hs-hninon-'	ne	kitkit	**aoti**-khw-a'
FUT-2SG.agt-buy-PFV	ART	chicken	**FZ.AL.DP.POSS**-food-NS
you will buy	ART	chicken	**their** food

'You'll buy the chickens' food.'

2.3.5 Nominalized attributes

There is no adjective category in Mohawk. Qualities are predicated with morphological verbs, complete with their pronominal prefixes. When these are used as referring expressions, they retain the number distinctions of verbal pronominal prefixes: singular/dual/plural for third person agent animates, and singular/duo-plural for third person patient animates.

(56) Attribute: Rita Konwatsi'tsaién:ni Phillips, speaker p.c.

O:nen	*kí:ken*	*ne*	*th**a**kowá:nen*
onen	kiken	ne	t-**ha**-kowan-en
now	this	ART	CSL-**M.SG.AGT**-be.large-ST
now	this	ART	**he** is bigger

'Now the **older** boy'

tahohteròn:ne'.
ta-ho-hteron-'ne-'
CSL.FAC-M.SG.PAT-be.afraid-AMB-ST
he is coming along afraid

'began to get scared as he went along.'

(57) Attribute: Rita Konwatsi'tsaién:ni Phillips, speaker p.c.

Iakèn:'ak   iahninhohó:ka'te'
iaken'a=k   i-a-hni-nhoh-oka't-e-'
barely=just   TRL-FAC-M.DU.AGT-door-reach-EP-PFV
just barely   they two reached the door there
'They just barely reached the door'

roti*kstèn:ha*   ***i****atathróna'.*
roti-ksten=ha   **i**-atat-hrona'
M.DP.PAT-be.old=DIM   **M.DU.AGT**-REFL-be.married
they are old   **they two** are married to each other
'of an old couple.'

Since the verb 'be married' occurs with agent prefixes, the husband and wife are referred to as dual in 'they two are married' = 'married couple', but since the verb 'be old' appears with patient prefixes, they are referred to just as duo-plural in 'they are old' = 'old people'.

2.4 Verbal number

Verbs may contain distributive suffixes that spread events or states over places. There is a special distributive suffix *-'s* for statives.

(58) Verbal distributives
  a. *Será:kew!*
    se-rakwe
    2SG.AGT-wipe
    'Wipe it!'

    *Serakew-**ánion**!*
    se-rakwe-**anion**
    2SG.AGT-wipe-**DISTR**
    'Wipe around **everywhere**! Dust **everywhere**!'

  b. *Satónhew!*
    s-atonhew
    2SG.AGT-sweep
    'Sweep!'

    *Satonhew**ánion**!*
    s-atonhew-**anion**
    2SG.AGT-sweep-**DISTR**
    'Sweep **around** (**all over** the house)!'

   c. *Wakkéhte'.*
      wak-keht-e'
      1SG.PAT-carry.on.back-ST
      'I am carrying it on my back'

      Wakkéhte-'s
      wak-keht-e'-'s
      1SG.PAT-carry.on.back-ST-ST.DISTR
      'I am carrying it **around** on my back'

They can imply that individual acts are carried out at various times. If I hid in various places, it might be inferred that I did not hide in them all at once.

    Verbal distributives can also describe actions or states distributed over multiple participants.

(59) Verbal distributives
   a. *Senóhare!*   'Wash it/them!'
     Senoharé-nion! 'Do the washing!' (wash **various** things)

   b. *Seksóhare!*   'Wash the dish!'
     Seksoharé-nion! 'Wash the dish**es**! (**each one**)'

   c. *Tstá'tha't!*   'Dry it/them!'
     Tsta'tha't-ánion! 'Dry **each** of them!'

   d. *Sà:swaht!*   'Turn the light off!'
     Sa'swaht-ánion! 'Turn the lights off! (**each individually**)'

   e. *Raenté:ri*   'He knows it or her'
     Raenté:ri-'s 'He knows **each one**'

   f. *Kowá:nen*   'It is big'
     Kowá:nen-'s '**Each one** is big'

Distributives do not specify plurality of arguments, but they can imply that an action is directed at objects, or a state spread over multiple individuals. They also suggest that the entities over which the action or property is distributed are various and assorted. The first verb in (60) could be used for buying one object or many. The second is more specific: it would not be used, for example, for buying a bouquet of flowers or a bag of apples.

(60) a. *Wa'khní:non'.*
     wa'-k-hninon-'
     FAC-1SG.AGT-buy-PFVL
     'I bought it/them.'

    b. *Wa'khninón**nion**-'*
       wa'-k-hninon-nion-'
       FAC-1SG.agt-buy-**DISTR**-PFV
       'I bought **various things**.'

There is also a prepronominal repetitive prefix on verbs that can indicate repetition of an action or return to an earlier state.

(61) Repetitives
    a. *Tsì:ron*   'Say it!'
       ***Sa**-tsì:ron* 'Say it **again**! **Re**peat it!'

    b. *Será:kew*   'Wipe it!'
       ***Sa**-será:kew* 'Wipe it **again**!'

    c. *Sá:rat*   'Lie down!'
       ***Sa**-sá:rat* 'Lie **back** down!'

    d. *Shahseró:ten*   'Turn the light on!'
       ***Sa**-shahseró:ten* 'Turn the light **back** on!'

Another prepronominal prefix, the duplicative (abbreviated DV, earlier termed the dualic), is derivational, appearing in lexicalized combinations with verb roots. In a few combinations a semantic element of duality is clear. It combines with the verb root *-ke* 'be multiple' to yield a stem meaning 'be two in number', as in (31) 'two skirts' seen earlier. It combines with the verb root *-iahse-* to yield a stem meaning 'be a pair', as in (35) 'a pair'. It combines with a verb root *-neken* to yield a stem meaning 'be side by side'. In other combinations an element of duality can be discerned, but somewhat less obviously, indicating a change of position or state. It appears with the stem *-ta'n* (*-t-a'n-* 'be.standing-INCHOATIVE) to yield verbs meaning 'stand up' and 'stop'. It does not, however, appear in verbs meaning 'sit down'. It appears in verbs meaning 'pick something up', but not in verbs meaning 'drop'. It appears in a verb meaning 'sing'; this is literally 'pick up one's song', based on the same verb stem 'pick something up' with an incorporated noun. In combination with the reflexive prefix it forms reciprocals.

    There is little evidence of verbal number distinctions within verb roots or stems, or of suppletive forms alternating with number of arguments.

3 Agreement and the syntax of number

There is no agreement in Mohawk. The pronominal prefixes on nouns and verbs identify their referents directly, rather than reflecting concord with any other elements in the sentence. The sentence in (62) is about just one family, but the pronominal prefixes on the verbs 'live' and 'have children' are duo-plural and plural.

(62) Direct reference: Rita Konwatsi'tsaién:ni Phillips, speaker p.c.
Karhá:kon   *t**hati**'terón:tahkwe'*   *iá:ken'*
ka-rh-akon   t-**hati**-'teront-ahkwe'   iaken'
N-forest-interior   CSL-**M.PL.AGT**-reside-PAST   HRS
in the forest   there **they** used to live   they say
'In the forest there used to live, they say,'

*kí:ken **ka**hwá:tsire'*.
kiken   **ka**-hwatsir-e-'
this   **N**-family-EP-NS
this   family
'this **family**.'

*Tóhka' ni**hoti**wí:raien'.*
tohka'   ni-**hoti**-wir-a-ien-'
several   PRT-**M.DP.PAT**-child-LK-have-ST
several   **they** children have
'**They** had several children.'

There is scant evidence in Mohawk of a determiner phrase or noun phrase. There is just one article, *ne* 'the aforementioned', which occurs directly before the nominal it modifies, but it does not distinguish number. As noted in section 2.2.3, the two demonstratives *kí:ken* 'this one, these' and *thí:ken* 'that one, those' are referring expressions on their own, though they may occur in apposition to other referring expressions. They do not necessarily occur adjacent to those expressions, however, or even in the same prosodic phrase.

(63) Demonstrative: Charlotte Kaherákwas Bush, speaker p.c.
Tóka   *ken*   *sheienté:ri*   ***kí:ken***   ...
tóka   ken   she-ienteri   **kí:ken**
maybe   Q   2SG>fi-know   **this**
maybe   ?   you know her   **this one**
'Do you know **this person**,'

Ahkwesahshró:non'?   ...
ahkwesahs=hronon'
Ahkwesahsne=resident
'this person from Ahkwesahsne?'

Shikenkwitè:ne	***kí:ken***	*wa'khé:ken'.*
shi-ka-ikwite=hne	**kí:ken**	wa'-khe-ken-'
COIN-N-spring=place	**this**	FAC-1SG>fi-see-PFV
in the spring	**her**	I saw her

'I saw **her** in the spring.'

In any case, the demonstratives do not distinguish gender or number, so they could not agree.

Similarly there are no adjectives. As seen in section 2.3.5, qualities are predicated with verbs. Such verbs may be used as referring expressions, as seen in (56) and (57), but they do not combine tightly with nouns as attributives to form phrases. They may of course incorporate noun stems, as in *katsi'noniowá:nen's* 'big bugs' seen earlier in (45), repeated here.

(45) ***ka**tsi'noniowá:nen's.*
ka-tsi'nonw-owan-en-'s
N-bug-be.large-ST-DISTR
'big bugs'

Otherwise, they may appear in apposition to another nominal expression, but if they do, there is no concord linking the two. It is of course not unlikely that they might show the same number distinction in their prefixes, since they would be identifying the same referent, but they also may not, depending on the possibilities afforded by the nouns or verbs in which they appear. As seen in examples (44) and (45) in section 2.3.2, it is not uncommon to see number mismatches between nominals and associated predicates when the nominal is a lexicalized form of a verb. Quantifiers can stand on their own, and may or may not be coreferential with a pronominal prefix or other element in a sentence. In (62) 'they had several children', 'children' is not an argument, but simply an incorporated noun qualifying a kind of having. 'Several' is understood as quantifying the children, but this is not agreement. The sentence would be just as grammatical without the incorporated noun: 'They had several'.

4 Semantics and discourse

Generic statements appear in a variety of forms, much as in English. Some contain a lexical nominal unmarked for gender or number, and a verb with third person plural pronominal prefix.

(64) Generic humans: Vina Loft, speaker
***Ron**ttókha'*	*thí:,*
ron-at-tok-ha'	thiken
M.PL.AGT-MID-notice-HAB	those

'They are smart,'

o'seron:ni.
o-a'ser-onni
N-axe-make
'white people.'

(65) Generic animals: Ima Johnsons, speaker

Kátke nón:	*ta**hón**tera'ne'*		*ne*	*è:rhar,*
katke	nonhwa'	t-a-**hon**-ate-ra'n-e-'	ne	erhar
when	time	DV-IRR-**M.PL.AGT**-REFL-meet-EP-PFV	ART	dog

'When dogs run into each other,'

nen	*kia'*	*tiótkon*	*óksa'ok*	*ia**hon**tkáhtho'* ...
onen	kia'	tiotkon	oksa'=ok	i-a-**hon**-at-kahtho-'
then	just	always	immediately	TRL-IRR-**M.PL.AGT**-MID-look-PFV

'they always immediately look ...'

Some generic statements contain a lexical nominal derived from a verb with a gender prefix but inert singular number marking, and a predicate that is a verb with plural pronominal prefix.

(66) Generic humans: Minnie Hill, speaker

Tóhsa'	***ra**'serón:ni*		*te**ietshi**'nikonhrhá:ren*
tohsa'	**ra**-a'ser-onni		te-**ietshi**-'nikonhr-haren
PROH	**M.SG.AGT**-axe-make		DV-**2PL>3PL**-mind-hang
don't	**he** axe makes		**you** bother with **them**

'Don't bother with white people.'

Some consist of a lexical nominal formed from a verb with a singular pronominal prefix, and a predicate that is a verb with a singular pronominal prefix.

(67) Generic humans: Wilfred Jaimison, speaker

***Ra**'serón:ni* ...
ra-a'ser-onni
M.SG.AGT-axe-make
he axe makes
'The white man'

***ra**nekwenhsáweron*
ra-nekwenhs-aweron
M.SG.AGT-blood-spill
he blood spills
'spills blood'

oh	naho:ten'	te**ho**tenhwentsio:ni	a**ha**ié:na'.
oh	naho'ten'	te-**ho**-aten-hwentsioni	a-**ha**-iena-'
what	thing	DV-**M.SG.PAT**-MID-want	IRR-**M.SG.AGT**-take-PFV
what	thing	**he** wants	**he** would take

'for whatever he wants to take.'

Neuter generics, like other neuters, show no number distinctions.

(68) Generic inanimates: John Maracle, speaker

Tsiohtsió:ri ratia'tónhkwahs
tsiohiori   rati-ia'ton-hkw-ahs
black.birch   M.PL.AGT-call-INS-HAB
'Black birch they call it.'

***Io**whá:rote'.*
io-whar-ot-e-'
N.PAT-fur-stand-EP-ST
'It has hair on it.'

Second person singular forms are also used in generic statements, much as in English.

(69) Second person: Minnie Hill, speaker
*Iáh ò:ni thaón:ton   a**hshe**hró:ri'   nahò:.*
iah ohni th-aa-w-aton   aa-**hshe**-hrori-'   naho'ten'
not also CONTR-IRR-N.AGT-be.possible IRR-**2SG>3pl**-tell-PFV anything
'You can't tell them anything.'

The third person feminine-indefinite pronominal prefixes are used for indefinite and generic persons, when sex is unknown or unspecified: 'one, people'. In this use it has no dual or plural forms.

(70) Indefinite gender: Mae Niionronhà:'a Montour, speaker p.c.
*Ónhka'k   wa**hshakó**:ken'   ta**iakaw**enonhátie'.*
onhka'=ok wa-**hshako**-ken-'   ta-**iakaw**-e-n-on-hatie'
who=only   FAC-**M.SG>FI**-see-PFV   CSL-**FI.PAT**-go-DIR-ST-PROG
someone   he saw **one**   **one** is coming along
'He saw someone coming.'

(71) Indefinite gender: Billy Kaientarónkwen Two Rivers, speaker p.c.
('It doesn't matter what I say.')

Ne:ne   *en**iako**thón:te'*   *ne'*   *ne:*
nene   en-**iako**-at-hont-e'   ne'   ne:
that.one   FUT-**FI.PAT**-MID-listen-ST   that   that
that one   **one** will be listening   that
'Whoever is listening,'

ne:ne: *ten**ie**ia'tó:re'ne'*   *nahò:ten'* *enkì:ron'.*
nene   t-en-**ie**-ia'tore'n-e-'   naho'ten' en-k-ihron-'
that.one DV-FUT-**FI.AGT**-decide-EP-PFV what   FUT-1SG.agt-say-PFV
that one **one** will decide   what   I will say
'will decide what I'm saying.'

Indefinite prefixes show up in large numbers of terms for tools and other equipment, terms coined as verbs describing their functions.

(72) Indefinite gender
ie*ksokewáhtha'*
ie-ks-okew-a-ht-ha'
FI.AGT-dish-wipe-LK-INS-HAB
'**one** dish wipes with it' = 'dishtowel'

This indefinite category has developed an additional use in the Northern Iroquoian languages, to show deference in referring to certain female persons, particularly grandmothers and mothers, but also many others. This is why it is now termed feminine-indefinite. In some of the languages, this is now the only gender category used to refer to female persons. In its indefinite use, the category is in a sense vague about number: 'one, people'. It its feminine use, it is singular: 'she, her'. When Mohawk speakers wish to refer to multiple female persons, each of which would be categorized as feminine-indefinite, the same forms are used as those classified as feminine-zoic. The dual and plural feminine prefixes are based on the feminine-zoic forms.

(73) Feminine prefixes
ka-   FEMININE-ZOIC.SG.AGENT   *ie-*   FEMININE-INDEFINITE.SG.AGENT
keni-   FEMININE.DU.AGENT
konti-   FEMININE.PL.AGENT

io-   FEMININE-ZOIC.SG.PATIENT   *iako-*   FEMININE-INDEFINITE.SG.PATIENT
ioti-   FEMININE.DUO-PLURAL

(The *-ni-* dual and *-ati-* plural elements of these pronominal prefixes are easily segmentable. The same elements appear in masculine prefixes: *ra-* M.SG.AGT, *hni-* M.DU.AGT, *rati-* M.PL.AGT, and *ro-* M.SG.PAT, *roti-* M.DP PAT.)

5 Conclusions

Mohawk shows clear singular, dual, and plural inflectional number distinctions, but the specific distinctions available differ across and within lexical categories. There are three basic word classes, defined by Iroquoianists in terms of their morphological structure: verbs, nouns, and particles. The largest number of distinctions are made within the verb and noun morphology. Number is distinguished in verbs in the pronominal prefixes referring to core arguments, and in nouns for possessors in possessive prefixes. The number of distinctions available depends on both gender and grammatical role. There are four genders, distinguished only in third person: masculine (for male persons and a few male animals), feminine-zoic (for some female persons and most animals), feminine-indefinite (for other female persons and unidentified persons), and neuter (for inanimates). For feminine-indefinites used for indefinites, there are only singular forms. For neuters, number is not specified.

On verbs, singular, dual, and plural are distinguished in grammatical agent prefixes (apart from feminine-indefinites and neuters), and in first and second person grammatical patient prefixes. Dual and plural are not distinguished in third person patient prefixes, nor in transitive combinations involving third persons. Nouns themselves are all grammatically neuter. But they may contain inalienable possessive prefixes showing the same distinctions as verbal agent prefixes, or alienable possessive prefixes showing the same distinctions as verbal patient prefixes, for possessors.

Independent particle pronouns are used only in certain situations for various kinds of contrast or when there is no verb. They do not distinguish grammatical role: the same forms refer to grammatical agents, patients, inalienable possessors, and alienable possessors. They also show fewer number distinctions than the verb and noun prefixes: only a singular/duo-plural distinction in third persons. Other morphological particles, determiners and indefinite and interrogative pronouns, do not distinguish gender.

Though the lexical categories are clearly identifiable from their morphological structures, these categories are not necessarily isomorphic with their syntactic functions. In particular, morphological verbs can serve not only as predicates and full sentences (complete with identification of core arguments in the pronominal prefixes), but also as referring expressions, often without further marking such as nominalizers or associated articles or demonstratives. Such verbal nominals may show the same number distinctions as their verbal sources. In some cases, however, once they have become nominalized, their prefixes can become inert, no longer shifting with number.

Where number distinctions are available, they are obligatory and straightforward: singulars always refer to just one, duals to just two, duo-plurals to two or more, and plurals to three or more. The same principles underlie categorization in pronominal prefixes on verbs, possessive prefixes on nouns, and particle pronouns,

even though the distinctions marked differ with person and grammatical role. Number distinctions do not, however, appear to have repercussions beyond the word. There is no evidence of agreement within either phrases or clauses. Each form refers directly, mirroring the number of its referent according to the distinctions available.

Overall, the basis for number marking is the same across lexical categories, directly specifying number and obligatory. The options available vary, however, with word class, gender, and grammatical role. Most significantly, the pronominal prefixes on verbs and possessive prefixes on nominals show more elaborate distinctions than independent pronouns.

Abbreviations used in glosses

AGT	grammatical agent
ART	article
AUG	augmentative
CMPL	completive
CONTR	contrastive
CSL	cislocative
DIM	diminutive
DISTR	distributive
DP	duo-plural
DV	duplicative
EP	epenthetic
EX	exclusive
FAC	factual
INCH	inchoative
INS	instrumental applicative
IRR	irrealis
MID	middle
NMLZ	nominalizer
NS	noun suffix
PAT	grammatical patient>
PFV	perfective
PL	plural
PRT	partitive
REP	repetitive
REV	reversive
ST	stative
TRL	translocative
Z	zoic

Reference

Mithun, Marianne. Forthcoming. The Manmade versus the Natural: Iroquoian Nouns in ka- and o-. In David Costa, Amy Dahlstrom & Lucy Thomason (eds.), Linguistics, Ethnohistory, and Reconstruction. East Lansing: Michigan State University Press.

VI Further perspectives on linguistic diversity

Viveka Velupillai
21 Contact languages: A survey

Abstract: This chapter discusses number in pidgins, pidgincreoles, creoles and mixed languages. It will show that it is more common for most of these languages to mark number in pronouns than in nouns. Clusivity, however, is not commonly marked for any of these languages. Pidgin languages tend to not have nominal plural marking, nor agreement marking in verbs or modifiers. The associative plural is not common for pidgins. Nominal number tends to be optional for pidgincreole and creole languages, typically marked with a plural word. The associative plural is common for both pidgincreoles and creoles. As for number agreement and syntax, adjectives tend to remain in their base form for both creoles and pidgincreoles, and verbs are typically not marked for number agreement. Determiners, however, do tend to show number agreement with their head. Coordinating constructions are not redundantly marked for number. Mixed languages tend to reflect the system of one of the input languages. Generic expressions tend to remain in the base form in all these languages.

Pidgins and pidgincreoles do not tend to mark politeness distinctions. Creoles, however, may optionally express politeness through the plural form. Politeness distinctions do not seem common in mixed languages.

1 Overview

This chapter gives a survey of number in contact languages. For the purpose of the discussion in this chapter contact languages are narrowly defined as those languages which have been classified as pidgins, extended pidgins (pidgincreoles), creoles or mixed languages. Other high contact varieties are thus beyond the scope of this chapter.

A **pidgin** language is here defined as "a structured language that emerges through the need of a communicative bridge between speakers of mutually unintelligible languages; it is usually learned as a second language and is typically used in specific situations or as a *lingua franca* across communities" (Velupillai 2015: 17). By lingua franca I am here referring to a language that is systematically used to enable communication between two groups of people who have different L1s and that is not native to either of the groups. For a useful overview of pidgin languages, see, for example, Parkvall (2017).

An **extended pidgin (pidgincreole)** is a contact language that started out as a pidgin but which has become a main means of inter-ethnic communication. It has typically become a main language for its community and is regularly employed in a large and expanding number of situations, in fact potentially in any situation.

This is in contrast to pidgin languages, which tend to be employed only in a limited amount of specific situations. Extended pidgins (pidgincreoles) may also become a mother tongue to some of its speakers. There has been considerable discussion about whether this expansion of domains, usage and users might make these languages more similar to creoles than to pidgins. However, the fact that they are for the most part additive languages (i.e. L2s) and not mother tongues for the majority of their speakers differentiate them from creoles. The term 'pidgincreole', first suggested by Philip Baker and widely established via Bakker (2008), is thus increasingly used, in an effort to show that these languages share affinities with both pidgins and creoles. In this chapter the term pidgincreole will henceforth be used.

A **creole** language is a natural language spoken as a mother tongue by an entire community that arose in situations of intense contact, where people of diverse ethno-cultural and linguistic backgrounds were brought together and formed distinct communities (Velupillai 2015: 43). They are full-fledged languages on a par with any other natural language.

Relevant for all three types of languages defined above is the distinction between Atlantic, Pacific and Indian Ocean contact languages. The Atlantic contact languages are those that share the Atlantic Ocean, that is, pidgin, pidgincreole and creole languages that are spoken on the eastern coasts of the American continent, on the Caribbean islands, and on the western coasts of the African continent. The Indian Ocean languages are those languages spoken in areas which all share the Indian Ocean and the Pacific languages are those spoken in or around the Pacific Ocean.

Relevant for them all are also the various types of input languages to their formation. A **substrate language** is usually described as the input language which has provided the structure for a contact language, but that can be misleading, as structural influence can come also from other input languages. More specifically it is (i) a language which is already established for a person and may influence the outcome of the acquisition of another language and (ii) thus on a more extended level is that language or variety which has influenced the structure or use of a more dominant language or variety of a speech community. A **lexifier language** is the language from which a contact language derives most of its lexicon. A **superstrate language** is a language or variety which has influenced a less dominant language or variety in a community. The lexifier and superstrate languages are often the same for a given contact language, but not always. An example of where they are not the same is for Sranan (Creole (English-based): Suriname), which has English as its lexifier language, but where Dutch has been the superstrate language since the 17th century.

Mixed languages are languages which came about through the fusion of two or a few identifiable source languages and thus take part of their linguistic system from one source and the other part from the other source language. They typically represent identity markers that emerged through expressive rather than communi-

cative needs, most commonly in situations of stable community bilingualism (Velupillai 2015: 69). There are essentially three types of mixed languages: a G-L language has the bulk of its grammar (both the structure and the forms of the markers) from one source language and the bulk of its lexicon from the other source language; an N-V language has the bulk of its noun phrase (both lexically and structurally) from one source language and the bulk of its verb phrase (both lexically and structurally) from the other source language; an F-S language has the bulk of its forms from one source language and the bulk of its structure from the other source language. In other words, in F-S languages all forms (lexical as well as grammatical) come from the same source, but the syntactic structure comes from another source. These classifications were initially suggested by Peter Bakker and have been widely adopted. See further Bakker (2020) and the references therein.

For more detailed definitions of pidgins, pidgincreoles, creoles and mixed languages, see Velupillai (2015).

This chapter is structured as follows: Section 2 gives a survey of pronominal, nominal and the associative plural in each of these types of contact languages, with a mention about clusivity and verbal number. Section 3 gives an overview of agreement and the syntax of number in contact languages. In Section 4 focus is given to the discourse function of number in contact languages, specifically politeness distinctions and genericity. Section 5 sums up the patterns found in pidgins, pidgincreoles, creoles and mixed languages respectively.

Each section discusses the subject at hand first for pidgins, then for creoles. Pidgincreoles are then discussed within the creole section. This will allow the reader to first get acquainted with both the patterns found in pidgin languages and the patterns found in creole languages, in order to then be able to ascertain which type of language the pidgincreoles show more affinities with. Finally mixed languages are given their own section. The patterns discussed for each type of language are based on the sample of pidgins, pidgincreoles, creoles and mixed languages respectively in Velupillai (2015). Details of each feature sample can be found online at https://benjamins.com/catalog/cll.48/additional (last access 7 December 2020). I should point out that I use the term **system** to refer to an organisation or paradigm while I use the term **form** to refer to the actual shape or structure of a given marker.

For each example the language name, type, major lexifier or input languages and approximate location have been given. This is also given the first time a language is mentioned, unless an example with the metadata is immediately following.

2 Pronominal, nominal, and verbal number

2.1 Generalities

Overall, pidgins, pidgincreoles, as well as creoles show an asymmetry in number marking between nouns and pronouns. It is generally more common for number to be expressed in the pronominal system than for nouns.

Pidgins do not tend to mark nominal plural, although, overall, there may be possibilities to mark number if needed (sometimes constrained by animacy in that only human nouns are marked for number). However, they do tend to differentiate between singular and plural in the pronominal systems, very often taking over the system of the lexifier.

Creoles also tend to display an asymmetry between nouns and pronouns but differ from pidgins in that most commonly all nouns can optionally be marked for plural by some means, without the animacy restriction found in pidgins. The plural marker is most commonly a specific plural marker or, in about the same amount of languages, the same form as the third person plural independent pronoun. The plural marking strategies in creoles tend to be innovated, i.e. do not tend to be the same as in the lexifier languages. The vast majority of creole languages differentiate for number in the pronominal system by means of cumulative forms indicating both person and number.

The pidgincreoles pattern like the creoles in that they tend to have optional nominal plural marking and have suppletive cumulative forms in the pronominal system indicating both person and number.

Mixed languages also tend to show a certain asymmetry between nominal and pronominal number. The nominal number marking tends to be taken over from one of the input languages: for G-L languages it tends to be the taken over from the language that provided most of the structure while for N-V languages it tends to be taken over from the language that provided most of the noun phrase structure. The pronominal systems, however, tend to show a certain amount of variation and may often represent a blend of the input languages, both in terms of form and in terms of exponence.

The category of numeral classifiers is very marginal for pidgins and creoles, and thus far not known to be relevant for pidgincreoles or mixed languages.

There is no evidence for verbal number as distinct from agreement in any of the languages discussed here.

These generalisations also widely apply to non-contact languages. This, in fact, forms a core part of one of the most central debates in contact linguistics, namely whether different types of contact languages are typologically distinct or not, i.e. if different types of contact languages are structurally different from other kinds of contact languages in a predictable way. The purpose of this chapter is not so much to settle that debate one way or another, but to present the typical patterns exhibit-

ed by the different types of contact languages discussed here. In other words, this chapter will not specifically focus on typologically rare constructions that can be found in various contact languages – which would not help solve the debate – but rather on the most common patterns found in the types of contact languages discussed. The reader may then form his or her own opinion as to whether these patterns set contact languages apart from non-contact languages.

2.2 Pronominal number

2.2.1 Pidgins

Two main strategies for pronominal number seem common in pidgins. The first is to express plural through suppletion, making the forms cumulative or portmanteau markers coding both person and number. Typically the forms are the same as or similar to those of the lexifier. The personal pronoun system in Fanakalo (Pidgin (Zulu-based): South Africa), for example, derives from Zulu, while the Lingua Franca (Pidgin (Romance-based): Mediterranean basin) system derives from Romance (Table 1).

The second main strategy is to have a mixed system in that there are suppletive forms for the plural but there might also be syncretism in some or all persons. Chinese Pidgin English (Pidgin (English-based): China), for example, has a system very similar to English, which is typologically unusual in having number syncretism in the second person, except that the dependent personal pronouns might allow syncretism also in the third person.[1] Mobilian Jargon (Pidgin (Muskogean-based): USA) has number suppletion only in the first person and syncretism in the second and third persons. In other words, number distinction is limited in both these pidgins (Table 2).

It should be noted that the possibility of marking 'more than one' somehow may exist. In Yimas-Arafundi Pidgin (Pidgin (Yimas-based): Papua New Guinea), for example, the base form (singular), which distinguishes three persons, tends to be used in non-singular contexts. If it is necessary to emphasise the dual or plural number of the referents, the non-singular markers *kundamwin* (for dual) or *asəŋ* or *manba* (for plural) are be used (Table 3).

---

1 Whether Chinese Pidgin English has simply "taken over" the number syncretism of the second person from the lexifier English or whether it is in fact a general tendency to syncretism (given that this person also shows number syncretism) can be debated. This is not a trivial point, given that one of the central debates in pidgin and creole studies is whether or not these languages do away with the typologically rare features of their input languages. In this case it could be argued that Chinese Pidgin English has kept a typologically rare feature of English, contrary to general expectations; alternatively it could be argued that it has simplified its system by removing number distinctions also in the third person.

Tab. 1: The personal pronoun systems of Lingua Franca (Anonymous 1830: *iv*) and Fanakalo (Mesthrie & Surek-Clark 2013: 36).

	Lingua Franca	**Fanakalo**	**Zulu**
1SG	*mi*	*mina*	*miná*
2SG	*ti*	*wena*	*wená*
3SG	*ell(u)o* (M)/*ella* (F)	*yena*	*yená* [noun class 1]
1PL	*voi*	*thina*	*thiná*
2PL	*noi*	*nina*	*niná*
3PL	*elli*	*bona*	*boná* [noun class 1]

Tab. 2: The personal pronoun systems of Chinese Pidgin English (Matthews & Li 2013: 208) and Mobilian Jargon (Drechsel 1997: 113–4).

	Chinese Pidgin English		**Mobilian Jargon**
	dependent	**independent**	
1SG	*my, I, me*	*my*	*(e)no*
2SG	*you*	*you*	*(e)ʃ(no)*
3SG	*he, she, it*	*he/him, she/her*	*(e)lap*
1PL	*we*	–	*poʃno*
2PL	*you*	*you*	*(e)ʃ(no)*
3PL	*he, they*	*he*	*(e)lap*

Tab. 3: The personal pronoun systems of Yimas-Arafundi Pidgin (Foley 2013: 107).

	singular	**dual**	**plural**
1	*ama*	*kapa (kundamwin)*	*paŋgət (asəŋ)*
2	*mi*	*mi kundamwin*	*mi asəŋ/manba*
3	*mən*	*mən kundamwin*	*mən asəŋ/manba*

These non-singular markers are the same as would optionally be used to mark number for nouns. Notice that only the first persons have suppletive forms and thus an exponence coding both person (1st) and number (singular/dual/plural), while the other forms are monoexponential (separative) where the first form codes the person and the second form codes the number.

Clusivity does not tend to be differentiated in pidgins.

2.2.2 Creoles

The overwhelming majority of creole languages express pronominal number through suppletion, with cumulative forms indicating both person and number. The

Tab. 4: The independent personal pronoun systems of Saramaccan (Aboh et al. 2013: 31), Gullah (Klein 2013: 142) and Pichi (Yakpo 2013: 197).

	SARAMACCAN	GULLAH	PICHI	SOURCE
1SG	*mí*	*a*	*à*	*mi* / *I* (English)
2SG	*jú*	*ya*	*yù*	*you* (English)
3SG	*hɛn*	*e*	*è*	*he* (English)
1PL	*ú*	*we*	*wì*	*us* / *we* (English)
2PL	*únu*	*oona*	*ùna, ùnu*	*unù* (Igbo)
3PL	*dé*	*dey*	*dὲn*	*they* (English)

Tab. 5: The independent personal pronoun systems of Tayo (Ehrhart & Revis 2013: 174) and Kriol (Schultze-Berndt et al. 2013: 143).

	TAYO	SOURCE	KRIOL	SOURCE
1SG	*mwa*	*moi* (French)	*mi*	*me* (English)
2SG	*twa*	*toi* (French)	*yu*	*you* (English)
3SG	*lia, lja*	*il y a* (French)	*im*	*him* (English)
1DU.INCL	*nunde*	*nous deux* (French)	*yunmi*	*you and me* (English)
1DU.EXCL	–		*min(du)bala*	*me two fellow* (English)
2DU	*vunde*	*vous deux* (French)	*yundubala*	*you two fellow* (English)
3DU	*lende*	*leur deux* (French)	*dubala*	*two fellow* (English)
1PL.INCL	*nu*	*nous* (French)	*as, minolabat*	*us* / *mi* NEG *all-about* (English)
1PL.EXCL	–		*mibala, mela(bat)*	*mi all-about* (English)
2PL	*uso*	*vous autres* (French)	*yubala, yumob*	*you all-about* / *you mob* (English)
3PL	*sola, lesot*	*les autres-là* (French)	*olabat, dem*	*all-about* / *them* (English)

forms tend to derive from the lexifiers. However, the systems may differ from those of the lexifiers. For example, almost all creoles differentiate between the second singular and plural, meaning that the rare second person syncretism found in English has not carried over to the creoles. Instead, differentiation is made by various means. In the English-lexified Atlantic creoles it is very common for the second person singular form to be English derived (< *you*) while the second person plural form is derived from the Igbo *unù* '2PL', as can be seen in Saramaccan (Creole (English-based): Suriname), Gullah (Creole (English-based): USA) and Pichi (Creole (English-based): Bioko, Equatorial Guinea) (Table 4 above).

The creoles in the South Pacific area, such as Tayo (Creole (French-based): Saint-Louis, New Caledonia) and Kriol (Creole (English-based): Australia), have adopted the number systems of the substrates, thus also differentiating dual number.

Clusivity differentiation is rare in creole languages, but is found in Tayo, Kriol (Table 5 above), Sranan and Zamboanga Chabacano.

Tab. 6: The personal pronoun systems of Solomon Islands Pijin (Velupillai 2015: 421).

	SINGULAR	DUAL	TRIAL	PLURAL
1.INCL	*mi*	*iumitufala*	*iumitrifala*	*iumi*
1.EXCL	–	*mitufala*	*mitrifala*	*mifala*
2	*iu*	*iutufala*	*iutrifala*	*iufala*
3	*hem*	*tufala*	*trifala*	*olketa*

The pidgincreoles behave more like the creoles languages than the pidgin languages in my database. Thus all but Ghanaian Pidgin English distinguish between second person singular and plural. The languages spoken in the Oceanic area have adopted the dual, and in the case of Solomon Islands Pijin (Pidgincreole (English-based): Solomon Islands), also the trial numbers (Tab. 6).

The source of the Solomon Islands Pijin pronouns largely matches that of the pronouns in Kriol (Table 5): the singular forms are derived from English *me*, *you*, *him*; the dual forms from *you mi two fellow*, *mi two fellow*, *you two fellow* and *two fellow*; the trial forms from *you mi three fellow*, *mi three fellow*, *you three fellow* and *three fellow*; and the plural forms from *you mi*, *mi fellow*, *you fellow* and *all together*.

Clusivity differentiation has so far only been found in the Pacific pidgincreoles.

2.2.3 Mixed languages

Mixed languages show considerable variation in their pronominal systems, and very often the systems represent a blend of the two (or handful of) input languages. By far the most common is to express number through suppletion, giving the forms cumulative exponence. In Shelta, an English/Irish G-L Mixed Language spoken in Ireland, all forms are derived from English except the first and second singular, which are from Irish. The forms *dʲil* 'self' and the emphatic -*ʃa* are appended:

Tab. 7: The personal pronoun systems of Shelta (Velupillai 2015: 382).

	SINGULAR	PLURAL
1	*mwi:lʃa*	*our*
2	*tu:/hu:/di:lʃa*	*your*
3.M	*his dʲil*	*their dʲils*
3.F	*her dʲil*	

The Shelta first person singular is thus derived from Irish 1SG *mé* + *dʲil* 'self' + *ʃa* 'EMPHatic'; the second person singular is derived from Irish 2SG *tú* + *dʲil* 'self' + *ʃa* 'EMPHatic'. In Michif, a French/Cree N-V Mixed Language spoken in Canada, num-

Tab. 8: The personal pronoun systems of Michif (Velupillai 2015: 89).

	SINGULAR	PLURAL
1-EXCL	-	*nɪjana:n*
1-INCL	*nɪja*	*kɪjana:n*
2	*kɪja*	*kɪjawa:w*
3	*wɪja*	*wɪjawa:w*

Tab. 9: The personal pronoun systems of Gurindji Kriol (Meakins 2013: 133).

	SINGULAR	PLURAL
1-EXCL	*ngayu*	*ngaliwa*
1-INCL	*ngali*	*ngantipa*
2	*nyuntu*	*nyurru(lu)*
3	*nyantu*	*nyarru(lu)*

ber is expressed through affixation, making the plural forms separative in exponence. However, the pronominal forms, all derived from Cree, are only used for emphasis, as arguments are marked through verbal agreement (Table 8).

Clusivity differentiation is rare for mixed languages but is attested in the Gurindji Kriol (G-L Mixed Language (Gurindji/Kriol): Australia) emphatic pronouns, all of which derive from Gurindji (Table 9).

2.2.4 Summary

In sum pidgins do tend to differentiate between singular and plural in their pronominal systems, either through suppletion or portmanteau forms for a plural, or a mixed system with syncretism for some persons. Creoles and pidgincreoles typically express pronominal plural through suppletion. Mixed languages show a fair amount of variation and tend to exhibit a blend of the input language systems. Most common is that they express number through suppletion.

Tab. 10: Comparison of the typical patterns found in the different types of languages.

	PIDGINS	PIDGINCREOLES	CREOLES	MIXED LANGUAGES
PRONOMINAL PLURAL	no	yes	yes	yes
CLUSIVITY	no	no	no	no

2.3 Nominal number

2.3.1 Pidgins

It is most common for pidgins not to mark nominal plural (Velupillai 2015: 355f), meaning that the base form is used for any number, as in Romanian Pidgin Arabic:

(1) Romanian Pidgin Arabic (Pidgin (Arabic-based): Iraq)
     sadiq la   ani work la   rig
     friend PREP 1SG work PREP rig
     'My friend works on the oil rig.' / 'My friends work on the oil rig.'
     (Avram 2010: 22)

Here the reading of *sadiq* 'friend' can be either singular or plural, depending on the context. In Turku number was optionally marked with the suffix *-in/-en* or fossilized non-linear[2] forms:

(2) Turku (Pidgin (Arabic-based): Chad)
     a. *kam   inte fákar pfil(-in)   fi   gidam=na?*
        how.many 2SG think elephant(-PL) EXIST in.front=1PL
        'How many elephants do you think there are ahead of us?'

     b. *batán jijere   ma béji fi híle   íntekum kútulu yal   wála*
        again smallpox NEG come to village 2PL   kill   child.PL or
        besáo le=ku amyán
        make to=2PL blind
        'Never again will smallpox come to your village, kill your children or make you blind.'
        (Tosco & Owens 1993: 197, 200f)

A few pidgins have an animacy constraint in that while plural is generally not marked, it may optionally be expressed for human nouns, but only for human nouns, as in Chinese Pidgin Russian:

(3) Chinese Pidgin Russian (Pidgin (Russian-based): Russia)
     Tiper' ibəŋka   isio zenʃinə isio kərasivə.
     now   Japanese PL woman PL beautiful
     'Japanese women are beautiful.'
     (Perekhvalskaya 2008: 270)

---

2 A morphological marker which involves some kind of modification of its host. See further Velupillai (2012: 98–100).

Here the separative isolating form[3] *isio* (derived from Russian 'also, yet') has grammaticalized into an optional plural marker.

Numeral classifiers are typically not relevant for pidgin languages, but can be found in Chinese Pidgin English (Matthews & Li 2013) and in Tây Bòi (Pidgin (French-based): Vietnam; Reinecke 1971), though it should be noted that their use is more marginal than in their respective input languages.

Similarly, the associative plural is not commonly found among pidgin languages, although Pidgin Hawaiian has a special associative plural marker:

(4) Pidgin Hawaiian (Pidgin (Hawaiian-based): Hawaiian Islands, US)
  Nuinui   pilikia mi ame mama ma.
  plenty.of trouble 1SG and mother APL
  'There will be a lot of trouble for me, my mother, and her friends.'
  (Roberts 2013)

2.3.2 Creoles

It is most common for creoles to have optional plural marking for all nouns by means of a separative isolating marker (Velupillai 2015: 356–358), such as in Haitian Creole (Creole (French-based): Haiti). It should be noted, however, that there is an ongoing debate about analyticity in creole languages: it could be that the orthographic conventions, often taken over from English, where numerous compounds are spelled as separate words, might obscure instances of affixation in creole languages.

(5) Haitian Creole (Creole (French-based): Haiti)
  Li   kouri dèyè makout yo.
  3SG run   after brute   DEF.3PL
  'He ran after the brutes.'
  (DeGraff 2007: 117)

The isolating cumulative marker for definiteness and plurality *yo* (from French *eux*, the third person plural form) is optional and postposed to the noun it marks.

In some creoles plural marking is obligatory, such as in Cassamancese Creole, where the separative plural marker -(*wu*)*s* (from Portuguese -(*o*)*s*) is suffixed to the noun:

---

3 By separative isolating marker I mean a marker that is isolating in its fusion (i.e. it does not fuse with the host) and separative in its exponence (i.e. it expresses only one grammatical category); see further Velupillai (2012: 96–106).

(6) Cassamancese Creole (Creole (Portuguese-based): Senegal)
Ña   karu ta   lebá miñjer-us kabra-s kamati-s ku   l ansol-us pa
1SG.poss car   HAB carry woman-PL goat-PL tomato-PL with sheet-PL to
fera.
market
'I usually take women, goats, tomatoes and sheets to the market in my car.'
(Biagui & Quint 2013)

It seems equally common for creoles to differentiate between the third person plural form and the isolating plural marker as to use the same form for both functions (Maurer & the APiCS Consortium 2013: Feature 25). Some languages (seven to be exact) have an overlap: there are two distinct forms for the two functions (third person plural pronoun vs. nominal plurality), but one of the forms is monofunctional (i.e. can be used to express only one of the functions), while the other form can be used to express both the third person plural pronoun and the nominal plural. For those creoles that do not have an isolating plural form the feature discussed above is not relevant.

Numeral classifiers are typically not common in creole languages, but can be found in Gullah (Klein 2013), Sri Lanka Portuguese (Creole (Portuguese-based): Sri Lanka; Smith 2013) and Ambon Malay (Creole (Malay-based): Indonesia; Paauw 2013).

It is rather less common for creole languages to have the associative plural than for non-creole languages (Velupillai 2012). Even so, about two thirds of the creole languages in APiCS do have the feature associative plural, which is most commonly identical in form to the additive plural marker, such as in Reunion Creole:

(7) Reunion Creole (Creole (French-based): Reunion Island)
bann Linéda   bann marmay
PL   Lineda   PL   child
'Lineda and her family'   'children'
(Staudacher 2004: 57; Bollée 2013)

In some creoles the associative marker is identical with the third person plural pronoun, which, as is the case with Hawai'i Creole English, is not necessarily the same marker as the nominative plural:

(8) Hawai'i Creole English (Creole (English-based): Hawaiian Islands)
maɪ   fadɛ   dɛm justu   go   [...] sɛʔ [...] tɚdonɛʔ
1SG.poss father APL PST.HAB ACT   set   turtle.net
'My father and his friends / those associated with him used to set turtle nets.'
(own fieldwork data)

Hawai'i Creole English almost always marks plurality through a plural suffix -*s*, whereas the associative plural is identical to the third person plural pronoun. Some creoles have a special associative marker which is different from both the plural marker and the third person plural pronoun, such as Guinea-Bissau Kriyol:

(9) Guinea-Bissau Kriyol (Creole (Portuguese-based): Guinea-Bissau)
   ba Djon
   APL John
   'John and his associates.'
   (Intumbo et al. 2013)

In Guinea-Bissau Kriyol the associative marker is *ba* (from Portuguese *banda* 'band, side, gang, edge' or *bando* 'flock, band, gang, group'; Peter Bakker, personal communication), which is different from the plural marker suffix -*s* and the third person plural form *elis* (which in turn is SG *eli* + PL -*s*).

    Pidgincreoles pattern like creoles in that it is most common for them to have optional plural marking. In Nagamese, for example, the separative plural affix -*khan* may optionally be used for any noun, and is also the plural marker for pronouns.

(10) Nagamese (Pidgincreole (Assamese-based): India)
   a. *moy-khan to dud no-kha-y*
      1-PL   FOC milk NEG-drink-PRS
      'We don't drink milk.'

   b. *bosa-khan ketia ahi-se?*
      child-PL   when come-PST
      'When did the children arrive?'
      (Bhattacharjya 2001: 227f)

Numeral classifiers have not been found in pidgincreoles.
    With respect to the associative plural, pidgincreoles pattern like creoles rather than like pidgins in that it is more common for them to have that category. Most commonly the associative plural is identical in form with the additive plural marker, such as in Sango:

(11) Sango (Pidgincreole (Ngbandi-based): Central African Republic)
   a. *a-Ngunzapa a-eke ga na ala ka*
      PL-Ngunzapa PM-COP come PREP 3PL there
      'Ngunzapa and the others are coming over there.'

   b. *mo ke wara a-tagba   a-dole   a-nyen'*
      2SG COP find PL-cob.antelope PL-elephant PL-what
      'You'll find [in that place] cob antelopes, elephants, whatever.'
      (Samarin 2013)

As shown in example (11a), the form *a-* is used as an associative plural ('Ngunzapa-and.the.others'), while in (11b) it is used as a plural number marker. In Bislama, however, the associative plural has the same form as the third person plural pronoun *olgeta*:

(12) Bislama (Pidgincreole (English-based): Vanuatu)
    hem i   stap wetem Sale olgeta
    3SG AGR stay with   Sale 3PL
    'He's staying with Sale and his family.'
    (Meyerhoff 2013)

2.3.3 Mixed languages

Mixed languages tend to reflect the system of one of the input languages, typically the language which has provided most of the grammatical structure, or to reflect a blend of the input languages. Bilingual Navajo, for example, has taken over the Najavo system of not marking plurality in the noun, while Ma'a/Mbugu has taken over the Bantu system of obligatory cumulative prefixes which code noun class as well as number.

(13) Bilingual Navajo (G-L Mixed Language (<u>Navajo</u>/English): US)[4]
    <u>*bi-'éé'*</u>   *change <u>íí-łééh</u>*
    3SG.poss-garment change 1DU/PL-make
    'We are changing her clothes.'
    (Schaengold 2004: 54f)

(14) Ma'a/Mbugu (G-L Mixed Language (<u>Bantu</u>/Cushitic): Tanzania)
    a. <u>*m-lagé*</u>   <u>*va-lagé*</u>
       NC1-woman NC2-woman
       'woman'   'women'

    b. <u>*ki-hlatú*</u>   <u>*vi-hlatú*</u>
       NC7-finger NC8-finger
       'finger'   'fingers'
       (Mous 2003: 164)

Numeral classifiers are thus far not known to be relevant for mixed languages.

---

4 Henceforth the examples for mixed languages are differentiated graphically between the source languages in that the source language that is underlined in the metadata is also graphically marked with underlining in the examples.

It seems somewhat more common for mixed languages to lack the associative plural than to have it, and it is fair to propose that this is due to the fact that the respective input languages lack the feature associative plural. Of those languages that do have the associative plural, it tends to have taken over the form from the language that has provided the bulk of the grammar (for G-L languages), the bulk of the noun phrase (for N-V languages) or the bulk of the forms (for F-S languages). Thus Ma'a/Mbugu expresses the associative plural with a noun class marker (as in Bantu), while Gurindji Kriol expresses it with suffixes expressing animacy vs. inanimacy (as in Gurindji) and Sri Lankan Malay expresses it with the third person plural pronoun (similar to Malay):

(15) Ma'a/Mbugu (G-L Mixed Language (Bantu/Cushitic): Tanzania)
      va-akína túri
      2-folk   Turi
      'Turi and his people.'
      (Mous 2013)

(16) Gurindji Kriol (N-V Mixed Language (Gurindji/Kriol): Australia)
      Weya   ngakparn-nyarrara   nyila-rra
      where frog-APL   that-PAUC
      'Where's the toy frogs and other animals?'
      (Meakins 2007: 388)

(17) Sri Lankan Malay (F-S Mixed Language (Malay/Tamil/Sinhala): Sri Lanka)[5]
      Miflal-derang pasar-na blakang ka a-main-ðuðuk   ambɛ a-ðuuðung
      Milfal-APL   shop-DAT back   in PRS-play-PROG COMP PRS-AUX
      'Miflal and his friends are playing behind the shop.'
      (Slomanson 2013)

Gurindji Kriol also has a dual associative *-kuwang*:

(18) Gurindji Kriol (N-V Mixed Language (Gurindji/Kriol): Australia)
      Nyawa Becky-kuwang-tu tubala   ged-im biskit   yet
      this   Becky-DU-ERG   3DU.SBJ get-TR biscuit still
      'This Becky and her mate are still getting some biscuits.'
      (Meakins 2013)

---

5 The lexicon and forms of Sri Lankan Malay all derive from Malay, but the structure almost entirely mirrors that of Tamil and Sinhala, to the extent that Sri Lankan Malay is not intelligible to other Malay speakers (see further Bakker 2006). For that reason there is no underlining of the data in examples of Sri Lankan Malay (as all forms would be underlined).

Tab. 11: Comparison of the *typical* patterns found in the different types of languages.

	PIDGINS	PIDGINCREOLES	CREOLES	MIXED LANGUAGES
NOMINAL PLURAL MARKING	no	optional	optional	(varied)
NUMERAL CLASSIFIERS	no	no	no	no
ASSOCIATIVE PLURAL	no	yes	yes	no

In example 18 the dual suffix *–kuwang* functions as a dual associative 'X and his/her mate'.

2.3.4 Summary

In sum, pidgins do not tend to mark nominal plural and do not have numeral classifiers. The associative plural is not common. Creoles and pidgincreoles tend to have optional plural marking, typically with a free form marker. The associative plural is quite common for both creole and pidgincreole languages, though not fully as common as for non-creoles. Neither tend to have numeral classifiers. Mixed languages tend to reflect the system of one of the input languages. The associative plural and numeral classifiers are not common, probably because they are not found in the input languages of the known mixed languages.

3 Agreement and the syntax of number

3.1 Pidgins

The general absence of plural marking in pidgins means that the syntax of number marking is not relevant. Nouns and their modifiers tend to remain in their base form, as do verbs. Pidgin Delaware is an example of a pidgin with no number marking at all, also not in the pronominal system:

(19) Pidgin Delaware (Pidgin (Unami-based): US)
  a. *ne olocko toon*
    1   hole   go
    'We run into holes.'
    (Goddard 1997: 58)

  b. *chijr paétton mítzi suvvijvan mvvs*
    2   give   eat   all   animal
    'You give food (to) all (the) animals.'
    (Goddard 1997: 59, 61)

c. *kee squa og   enychan hatah?*
   2   wife and child   have
   'Do you have (a) wife and children?'
   (Thomason 1980: 173)

d. *nirj r̃ɵe, jõni matta manúnckus*
   1   say 3   NEG   bad
   'We should say, he (is) not bad.'
   (Goddard 1997: 59)

As can be seen in example (19),[6] the pronominal forms code person only, and all nouns (*olocko* 'hole', *mvvs* 'animal', *squa* 'wife', *enychan* 'child') and verbs (*toon* 'go', *paétton* 'give', *mítzi* 'eat', *hatah* 'have', *r̃ɵe* 'say') are in their base form.

Optional modification with quantifiers, such as 'many' or 'all' does occur in pidgin languages, and these quantifiers might grammaticalize into free form plural markers, as, for example, happened in the pidgincreole Bislama:

(20) Bislama (Pidgincreole (English-based): Vanuatu)
      ol nes   oli   bisi tumas
      PL nurse 3.PL.AGR busy too.much
      'The nurses were all too busy.'
      (Meyerhoff 2013)

In Bislama the plural marker *ol* (which is identical with the determiner) ultimately derives from a form of 'all'. The agreement marker *oli* is a cumulative marker denoting third person plural agreement and is different from the third person plural pronoun *olgeta* 'they'. It is a verbal agreement marker and this example therefore also shows that the predicative adjective (*bisi* 'busy') inflects like a verb.

Numeral modification may also occur in pidgins, in which case the numeral is most commonly preposed to the noun, which typically remains in base form, such as in Singapore Bazaar Malay:

(21) Singapore Bazaar Malay (Pidgin (Malay-based): Singapore)
      Saya ada   dua anak, satu jantan, satu perempuan
      1SG   have two child one male   one female
      'I have two children, one son [and] one daughter.'
      (Aye 2005: 109)

---

6 The differences in form in example (19) between *ne* and *nirj* have to do with assimilations, while the differences in form between *chijr* and *kee* have to do with orthographic conventions.

3.2 Creoles

As mentioned above, the dominant strategy for most creole languages is to mark nominal number with a plural word (Velupillai 2015: 356–358).[7] The proportions of preposed and postposed plural words are almost the same (17/31 or 54.8 % of the languages have a prenominal plural word while 14/31 or 45.2 % have a postnominal plural word). However, there is an areal difference in that systems with postnominal plural marking cluster in the Atlantic region, while systems with prenominal plural marking are fairly evenly spread across the globe.

(22) Jamaican (Creole (English-based): Jamaica)
Di pus dem av   nof   pikni
DET puss PL   have many child
'The cats have many kittens.'
(Farquharson 2013)

(23) Diu Indo-Portuguese (Creole (Portuguese-based): India)
Es   tud ε   kaz   də tud pad
this PL COP.NPST house of PL   priest
'These are the houses of the priests.'
(Cardoso 2009: 175)

In example (23) the plural marker *tud* postmodifies the determiner *es* 'this' but pre-modifies the noun *pad* 'priest'. There is also a difference in whether a plural word is the only plural marking strategy or merely one of the existing strategies: it is more common that prenominal plural free form marking is the only existing strategy, whereas only three creoles in APiCS (Nengee, Negerhollands and Guyanese) have postnominal free form plural marking as the only strategy. However, even if Jamaican, for example, may mark plurality through reduplication, the most common strategy is to mark number with a postposed plural word (Farquharson 2013).

Adjectives and other modifiers tend to remain in the base form, as in Ambon Malay:

(24) Ambon Malay (Creole (Malay-based): Indonesia)
ana~ana kacil
PL~child small
'small children.'
(Paauw 2008: 413)

---

7 But see above regarding the debate on whether this is due to orthographic conventions or actually due to morphological isolation.

In Ambon Malay number may be expressed through reduplication of the noun, but the modifying adjective remains in the base form.

Determiners, however, may be differentiated for singular and plural, even if nouns often remain in the base form. Plural marked determiners may in fact be the only marker for plurality inside the NP, as in Trinidad English Creole and Nengee.

(25) Nengee (Creole (English-based): Suriname, French Guiana)
Den mama den e poli den pikin.
DET.PL mother they IPFV spoil DET.PL child
'The mothers are spoiling the children.'
(Migge 2013)

(26) Trinidad English Creole (Creole (English-based): Trinidad & Tobago)
dis book dat book dem book
DEM.SG.PROX book DEM.SG.DIST book DEM.DIST.PL book
'this book, that book, those books'
(Solomon 1993: 84)

While Trinidad English Creole may optionally mark plurality through a suffix *−s*, the more common strategy is to use a plural word which is identical with the distal demonstrative plural form. The plural word may variously be preposed or postposed to the noun. If a plural word is used, the noun remains in its base form.

Numerals most commonly precede the noun, which tends to remain in the base form, such as in Virgin Islands Dutch Creole:

(27) Virgin Islands Dutch Creole (Creole (Dutch-based): Virgin Islands)
ēn man a ha dri jungkin
INDF man PST have three son
'A man had three sons.'
(de Jong 1926: 15)

Verbs do not tend to show number agreement. Instead free form pronoun words are used to indicate the subject person, such as in Rabaul Creole German:

(28) Rabaul Creole German (Creole (German-based): Papua New Guinea)
alle klane mɛnʃ, di hɔlən dise buç
PL small person 3PL get DEM book
'The boys, they are getting those books.'
(Volker 1982: 30)

Coordinated NPS tend not to be redundantly marked for number:

(29) Papiamentu (Creole (Spanish/Portuguese-based): Aruba, Bonaire and Curaçao)
dama i   kabayero   nan
lady   and gentleman PL
'Ladies and gentlemen.'
(Kouwenberg 2013)

In example 29 the plural marker is used only once and has a scope of both coordinated NPs.

Pidgincreoles show a more mixed picture in that plural marking strategies may vary. Even so, the most common strategy is to have a plural word, which in the Pacific pidgincreoles tends to precede the noun. When a plural word is used in Nigerian Pidgin (Pidgincreole (English-based): Nigeria) and Cameroon Pidgin English, however, it tends to follow the noun.

(30) Tok Pisin (Pidgincreole (English-based): Papua New Guinea)
askim ol lapun
ask   PL old.person
'Ask older people.'
(Verhaar 1995: 347)

(31) Cameroon Pidgin English (Pidgincreole (English-based): Cameroon)
pikin dem di   pley futbol   fo stad
child PL   IPFV play football for stadium
'The children are playing football at the stadium.'
(Ayafor 2008: 444)

Modifying adjectives tend to remain in the base form, as with both pidgins and creoles. In other words, pidgincreoles behave like creoles: those which optionally mark number with a plural word often do so with a determiner, commonly the third person plural form of the pronoun (cf. example 20 above). The noun remains bare.

Numerals almost always precede the noun. It may be variable whether the noun is inflected for the plural or remains bare, as in Ghanaian Pidgin English:

(32) Ghanaian Pidgin English (Pidgincreole (English-based): Ghana)
fɔ   waif twɛnti pikin-s
four wife 20   child-PL
'four wives, twenty children.'
(Huber 2013)

Verbs are not inflected for number; as with creoles, free form pronoun words are used to express the (pronominal) subject.

(33) Sango (Pidgincreole (Ngbandi-based): Central African Republic)
ala gwe na   Bangui lakwe
3PL go   PREP Bangui always
'They frequently go to Bangui.'
(Samarin 2013)

Coordinated noun phrases may optionally lack redundant number marking (as in example 20 above), or may, as in Sango, have plural marking on each noun:

(34) Sango (Pidgincreole (Ngbandi-based): Central African Republic)
a-mapa   a-nyama a-nyen'
PL-bread PL-meat PL-what
'bread, meat, whatever...'
(Samarin 2013)

Pidgincreoles thus pattern more closely to creoles than to pidgins in that number marking occurs more systematically.

3.3 Mixed languages

Mixed languages vary considerably with respect to the expression of plural meaning (cf. above). Very commonly the plural marker is a concatenative suffix. Michif, however, has most of its nouns from French, and they take the invariant prenominal article *lii* (< French *les*):[8]

(35) Michif (N-V Mixed Language (French/Cree): Canada)
Kaa-oshihtaa-chik   lii   maenzon lii   rosh kii-aapachih-eewak.
REL-make.INAN-3PL ART.PL house   ART.PL rock PST-use.ANIM-3PL
'When they made houses, they used rocks. OR: The houses that they made, they made of rocks.'
(Fleury 2007)

---

8 Notice that this is in fact nearly the same form as the French one, despite the orthographical differences, since in spoken French the <*-s*> is often left unpronounced.

Gurindji Kriol also has a paucal marker which appears as a suffix on the noun:

(36) Gurindji Kriol (N-V Mixed Language (<u>Gurindji</u>/Kriol): Australia)
 Wi garra ged-im <u>wumara</u>-<u>walija</u>.
 1PL.SBJ POT get-TR rock-PAUC
 'We'll get some rocks.'
 (Meakins 2013)

Modifiers do not necessarily inflect even if nouns are overtly marked for inflection. Thus Gurindji Kriol adjectives remain invariant, even though the noun may be inflected for number, whereas Ma'a/Mbugu modifiers, including numerals, inflect like nouns in that they take the relevant noun class marker:

(37) Ma'a/Mbugu (G-L Mixed Language (<u>Bantu</u>/Cushitic): Tanzania)
 a. <u>*vi*</u>-*aghú* <u>*vyó*</u> <u>*ni'á*</u> *ní* <u>*vi*</u>-*kuhló*
  8-food 8.that 1SG.eat is 8-nice
  'The food that I eat is nice.'

 b. *vigí* <u>*vi*</u>-*nú* <u>*ví*</u>-*áta* *i?í* *ní* <u>*vi*</u>-*ghó*
  things 8-two 8-be here is 8-my
  'The two things here are mine.'
  (Mous 2003: 205)

In example 37 the modifiers *vyó* 'that', *vikuhló* 'nice', *vinú* 'two', *viáta* 'be' and *vighó* 'my' are all marked for the noun class 8 which is semantically plural. Similarly, in Caló, adjectives agree with the noun in gender and number and in Media Lengua the determiner agrees with the head noun in number, as in Spanish.

(38) Caló (G-L Mixed Language (<u>Spanish</u>/Romani): Spain)
 <u>*sin*</u>-*aban* <u>*hambré*</u>-<u>*s*</u> *baribú* *latʃo*-<u>*s*</u>
 be-3PL.PST.IPFV person-PL much good-PL
 'They were very good people.'
 (Quindale 1867: 37)

(39) Media Lengua (G-L Mixed Language (<u>Spanish</u>/Quechua): Ecuador)
 <u>*isi*</u>-*guna* *mo* <u>*muxer*</u>-*pu* <u>*hermana*</u>-*guna*
 that-PL 1SG-poss woman-GEN sister-PL
 'Those are my wife's sisters.'
 (Muysken 2013: 145)

Numerals are typically not marked for plural, except that they behave as any other modifier in Ma'a/Mbugu (cf above).

The verbal morphology tends to follow that of the input language that has provided the bulk of the grammar (for G-L languages), the bulk of the VP (for N-V languages) or the structure (for F-S languages). Thus in Angloromani the verbs are not marked for number, similar to English verb morphology:

(40) Angloromani (G-L Mixed Language (<u>English</u>/Romani): UK)
    a. *mandi pukkere-<u>d</u> <u>the</u> rakli*
      1SG   say-PST   the woman
      'I told the woman.'

    b. *lesti=<u>s</u>   savv-<u>ing</u>   <u>at</u> mandi*
      3SG.m=prs laugh-PROG at 1SG
      'He's laughing at me.'
      (Matras 2010: 121f)

In Angloromani the verb is inflected for tense and aspect like in English, but the subject is expressed by free pronominal forms (as in English, but not with the English forms), which are cumulative forms denoting both person and number. Media Lengua, on the other hand, has portmanteau verbal affixes denoting person and number, like Quechua. Similarly, the Mednyj Aleut verbal system resembles that of Russian:

(41) Media Lengua (G-L Mixed Language (<u>Spanish</u>/Quechua): Ecuador)
    <u>entre</u>   <u>kwatro</u>-guna <u>i</u>-nchi
    among four-2SG.prs go-1PL
    'We go among the four of us.'
    (Muysken 1981: 63)

(42) Mednyj Aleut (N-V Mixed Language (<u>Aleut</u>/Russian): Commander Islands)
    taana-x̂   <u>ni</u>-buud-<u>ish</u>   ukuu-<u>t</u>'
    land-SG.ABS NEG-will-2SG.prs see-INF
    'You won't see the land.'
    (Sekerina 1994: 25)

As can be seen in examples 41 and 42, Media Lengua and Mednyj Aleut have verbal portmanteau suffixes expressing both person and number of the subject.

Coordinating constructions tend to resemble the structure of the input language that has provided the bulk of the grammar or the bulk of the NP to the language.

3.4 Summary

In sum, pidgins do not tend to have agreement or number marking. Verbal number is not expressed, and nouns and their modifiers tend to be in their base form. Cre-

Tab. 12: Comparison of the typical patterns found in the different types of languages.

	PIDGINS	PIDGIN-CREOLES	CREOLES	MIXED LANGUAGES
NOMINAL PLURAL MARKING	no	plural word	plural word	(as G or N input lg)
NOMINAL NUMBER AGREEMENT: MODIFIERS	no	no	no	(as G or N input lg)
NOMINAL NUMBER AGREEMENT: DETERMINERS	no	yes	yes	(as G or N input lg)
VERBAL NUMBER AGREEMENT	no	no	no	(as G or V input lg)
COORDINATING CONSTRUCTIONS REDUNDANCY	no	no	no	(as G or N input lg)

oles and pidgincreoles tend to mark non-singularity with a plural word. Adjectives tend to remain in their base form, while determiners tend to show number agreement with their head. Verbs do not tend to show number agreement. Coordinating constructions do not tend to be redundantly marked for number. Mixed languages tend to exhibit the pattern of that input language which has contributed the structure of the nominal or verbal phrases. In other words, G-L mixed languages tend to behave like the input language that provided the grammar with respect to agreement and the syntax of number marking, while N-V mixed languages tend to behave like the N-input language with respect to agreement and syntax of the noun phrases and like the V-input language with respect to agreement and syntax of the verb phrases. Coordinating constructions tend to resemble the input language which provided the bulk of the grammar or of the NP to the language.

4 Semantics and Discourse

4.1 Pidgins

Politeness distinctions are usually not found in pidgins (Velupillai 2015: 509), though in Tây Bòi the second person plural pronoun functioned as a second singular addressee honorific (as in French), as it did in Chinese Pidgin Russian:

(43) Chinese Pidgin Russian (Pidgin (Russian-based): Russia)
Kapitana waʃa ni   kasajs'a
captain   2PL   NEG concern
'It is none of your business, sir!'
(Jabłońska 1957: 167)

Singapore Bazaar Malay (Pidgin (Malay-based): Singapore) has the suppletive forms *awak*/*engkau* '2SG (polite but distant)' versus *lu*/*lia* '2SG (rude but intimate)' (Aye 2005:74), forms which are typical for all Malay-based contact languages.

Generic expressions tend to be expressed with the base form of the noun, as in Pidgin Delaware, although in Fanakalo it can be expressed either with a singular noun phrase and a definite article, or with a plural noun phrase and a definite article:

(44) Pidgin Delaware (Pidgin (Unami-based): North America)
　　　 ne olocko toon
　　　 1　 HOLE　 GO
　　　 'We run into holes.'
　　　 (Goddard 1997: 58)

(45) Fanakalo (Pidgin (Zulu-based): South Africa)
　　 a. *Lo　　　 pikanin yena hayi thand-a lo　　　　 pelepele*
　　　　 DEF.ART child　　 it　　 NEG like-V　 DEF.ART pepper
　　　　 'Children don't like pepper. / The child doesn't like the pepper.'

　　 b. *Lo　　　 bongolo yena balek-a kakhulu*
　　　　 DEF.ART donkey　 it　　 run-PRS greatly
　　　　 'Donkeys run greatly. / The donkey runs greatly. / A donkey runs greatly.'
　　　　 (Mesthrie 2013)

Given that plurality does not tend to be expressed, genericity, countability and other such common pragmatic functions of number are typically inferred through context using the base forms.

4.2 Creoles

It is more common for creole languages to lack politeness distinctions (Velupillai 2015: 511). However, about two fifths of the creole languages in my database do have some kind of politeness strategy, i.e. they have the possibility of expressing politeness by some means, even though they do not have obligatory politeness forms that stand in a paradigmatic relationship with informal forms. Of the languages that do have some kind of politeness strategy, it is most common with a binary politeness distinction, where the second person plural pronoun functions as the singular polite form, such as in Nengee (where the familiar form would be the 2SG *i* or *yu*):

(46) Nengee (Creole (English-based): Suriname, French Guyana)
gaanman,   u mu yeepi den   sama ya
paramount.chief 2PL must help DET.PL person DEM
'Chief, you should help these people.'
(Migge 2013b: feature 18)

Three creole languages, Korlai (Creole (Portuguese-India): South Africa), Cavite Chabacano (Creole (Spanish-based): Philippines) and Zamboanga Chabacano (Creole (Spanish-based): Philippines), have multiple politeness distinctions. In Zamboanga Chabacano, for example, *(e)bó(s)* or *tu* are the second person singular familiar forms, while *ustéd* is the polite singular form, and *kamó* is the second person plural form, while *ustédes* is the polite plural form (Steinkrüger 2013). A handful of the languages use pronoun avoidance (a politeness strategy where the speaker avoids using the pronoun when referring to someone, but uses some kind of NP) as their politeness strategy, such as Angolar:

(47) Angolar (Creole (Portuguese-based): São Tomé and Príncipe)
o   ma m bi kai Kompa [...] a   Kompa tha pê
hour REL 1SG come house friend [...] where friend be put
When I came to your house, [...] where were you?
(Maurer 1995: 62)

Generic nouns most commonly remain in the bare form, such as in Seychelles Creole:

(48) Seychelles Creole (Creole (French-based): Seychelles)
Lion i   manz gazel
lion PM eat   gazelle
'Lions eat gazelles.'
(Michaelis & Rosalie 2013)

Some languages may allow both kinds of strategies, such as Hawai'i Creole English:

(49) Hawai'i Creole English (Creole (English-based): Hawaiian Islands)
a. *da mɛnehuni-s   kam aʊt naɪtaɪm*
   DEF menehune-PL come out at.night
   'Menehunes ('little people') come out at night.'

b. *æs   wad dɛ hawaɪɛn dan*
   that.is what ART Hawaiian do.PST
   'That's what the Hawaiians did.' (own fieldwork data)

Pidgincreoles pattern more like pidgins than like creoles with respect to politeness in that it is more common for them to lack politeness distinctions, although Sango has a binary distinction:

(50) Sango (Pidgincreole (Ngbandi-based): Central African Republic)
mama   ala de   ti lango ape
mother 2PL continue of sleep NEG
'Mother, haven't you fallen asleep yet?!'
(Samarin 2013)

In Sango the second person plural pronoun is used as a polite form not only to a single addressee but also as a honorific when referring to a single third person. Bislama uses the second person plural form *yufala* as a polite address in some traditional respect kinship relationships (Meyerhoff 2013).

Generic noun phrases typically remain unmarked, as in Nagamese:

(51) Nagamese (Pidgincreole (Assamese-based): India)
kuta to   kotha-man-ia   jontu   ekta ase
dog FOC order-obey-NMLZ animal one be
'The dog is an obedient animal.' / 'Dogs are obedient animals.'
(Bhattacharjya 2001: 221)

Tok Pisin, however, expresses genericity through plural marking:

(52) Tok Pisin (Pidgincreole (English-based): Papua New Guinea)
Ol sikau   i   sae slip   antap long diwai
PL tree.kangaroo PL HAB sleep on.top PREP tree
'Tree kangaroos sleep in trees.'
(Smith & Siegel 2013)

In Nigerian Pidgin and Bislama both strategies are equally common.

4.3 Mixed languages

Politeness distinctions are not generally found in mixed languages, with only Sri Lankan Malay displaying a binary distinction in the pronominal system:

(53) Sri Lankan Malay (F-S Mixed Language (Malay/Tamil/Sinhala): Sri Lanka)
  a. *Lu   e-biilang (aða) attu buttul.*
     2SG.fam ASP-say AUX   one correct
     'What you have said is correct.'

b. *Lorang e-biilang (aða) attu buttul.*
   2SG.pol ASP-say AUX   one correct
   'What you (polite) have said is correct.'
   (Slomanson 2013)[9]

Little information is available on genericity in mixed languages, but it seems as if the strategy for expressing generic nouns is largely taken over from the input language that provided most of the structure. This would therefore follow the typical pattern for mixed languages. Thus Shelta, which has most of its grammar from English, and Caló, which has most of its grammar from Spanish, expresses genericity with the plural form:

(54) Shelta (G-L Mixed Language (<u>English</u>/Irish): Ireland)
      gre:di-<u>s</u>   ni:dʲa-<u>s</u>   ri:lu:
      make-3SG.prs person-PL mad
      '(It) makes people mad.'
      (Macalister 1937: 188)

(55) Caló (G-L Mixed Language (<u>Spanish</u>/Romani): Spain)
      sin<u>an</u>   ʧorrí-<u>as</u> <u>as</u>   kiribí-<u>s</u>   pur   pen-<u>an</u>   <u>as</u>
      be-3PL.PRS evil-PL   DEF.F.PL midwife-PL when speak-3PL.PRS DEF.F.PL
      ʧaʧumí-<u>s</u>
      truth-PL
      'The midwives are bad when they speak the truth.'
      (Quindale 1867: 22)

Similarly, Sri Lankan Malay has taken over the system of Malay and expresses genericity with the bare singular form:

(56) Sri Lankan Malay (F-S Mixed Language (Malay/Tamil/Sinhala): Sri Lanka)
      Kə-sattu kali-na   blakang kə-ðua   kali mana-waktu-le   a-datang
      ORD-one time-DAT after   ORD-two time where-time-COM PRS-come
      'The second time always follows the first time.'
      (Slomanson 2013)

⸺⸺⸺⸺

9 The polite form can be parsed as *lu* '2SG' + *orang* 'people', which in effect means that this is another instance of a plural form functioning as a polite form; note, however, that it here has acquired a paradigmatic function.

4.4 Summary

In sum, pidgins do not tend to have politeness distinctions. Generic expressions tend to be in the base form. Creoles may optionally express politeness through the plural form. Generic expressions tend to remain in the base form, as in pidgins. Politeness distinctions do not seem common in mixed languages. There is too little information available on genericity to draw any conclusions on potential patterns.

Tab. 13: Comparison of the *typical* patterns found in the different types of languages.

	PIDGINS	PIDGINCREOLES	CREOLES	MIXED LANGUAGES
POLITENESS DISTINCTIONS	no	(optional)	(optional)	no
GENERICITY	base form	base form	base form	(data N/A)

5 Conclusions

This chapter has discussed the category of number in contact languages. For the purpose of this chapter contact languages are narrowly defined as pidgin, pidgin-creole and creole languages, as well as mixed languages.

With regard to pidgins there is a rather noticeable morphological and semantic difference with respect to nominal and pronominal number. Most pidgins have a number distinction in the pronominal system, most commonly by means of suppletion (with or without person syncretism). The plural forms in pidgin pronominal plurals tend to be innovated rather than taken over by the lexifiers. By contrast, nouns are typically not marked for the plural. Numeral quantification tends to combine with the bare form of the noun. The associative plural is rare for pidgin languages. Pidgin languages typically do not have politeness distinctions in their pronominal systems and generic nouns are typically expressed with the noun in the base form. Number is thus a marginal category in pidgin languages, confined primarily to the pronominal system, but either optional or not relevant for nominal and verbal morphology and syntax. Overall it also seems to be a marginal category for the discourse structure of pidgin languages.

This differs from creole languages, where the category of number has a much more prominent role. The overwhelming majority of creole languages have number distinctions in the pronominal system. Interestingly, the English-lexified Atlantic creoles tend to avoid the number syncretism of the second person found in English. In other words, most of these Atlantic creoles differentiate between the second person singular and plural, as opposed to their lexifier. Some creoles in the Pacific area have adopted the dual number of the areal substrates. There thus seems to be a

drive for specificity in pronominal number distinction in these languages, irrespective of whether the lexifier has that or not.

While nominal number is largely irrelevant for pidgins, it is less so for creoles. Most creoles have optional plural marking by means of a separative isolating marker. This marker is rarely taken over from the lexifier, but is an innovated form. It is roughly equally common for this marker to be a plural word identical with the third person pronoun as it is for the plural word to be differentiated from it. It is also roughly equally common for creoles to have the plural word preposed to the noun as it is for them to have it postposed. Numeral classifiers are typically not attested in creole languages.

A further contrast between pidgin and creole languages is that while the associative plural is rare for pidgins, the majority of creole languages do have this category. Very often the associative plural form is identical to the additive plural form, though it may also be that the associative plural is expressed with the third person plural pronoun. In Guinea-Bissau Kriyol the associative plural form is different from both the additive plural and the third person plural pronominal form.

Creoles typically do not show number agreement inside the NP: adjectives and other modifiers tend to remain in the base form. Determiners, on the other hand, do tend to be marked for number but then the noun tends to remain in the base form so that number is only indicated on the determiner in the DP. Numerals tend to precede the noun, and as with determiner constructions, the noun tends to remain in the base form. Verbs are typically not marked for number but the VP is commonly number specified through the free pronominal forms.

Contrary to pidgin languages, a fair amount of creole languages have politeness distinctions in their pronominal systems. Most commonly this distinction is binary and the polite singular addressee form is expressed with the second person plural pronoun, though a few creole languages make use of pronoun avoidance. In creoles, genericity is typically expressed with the bare form of the noun.

Number is thus a much more central category for creole languages, morphologically, syntactically and semantically as well as in the discourse, than it is for pidgin languages.

Overall, pidgincreoles pattern like creoles with respect to the category of number. Pronominal number is always expressed and in most pidgincreoles the second person singular and plural pronoun is differentiated.

Similar to creoles, pidgincreoles tend to have optional plural marking, commonly, but by no means always, by means of a plural word which may be identical to the third person plural pronoun. Number is thus, like with creoles, indicated only on the determiner in a DP, while the noun remains in its bare form. Similar to creoles it is more common for pidgincreoles to have the associative plural, which most commonly is identical with the additive plural marker, though it occasionally is the third person plural pronoun that is used for this function. Numeral classifiers have not been found in pidgincreole languages. Modifying adjectives are not marked for

Tab. 14: Comparison of the majority patterns found in the majority of the different types of languages.

	PIDGINS	PIDGIN-CREOLES	CREOLES	MIXED LANGUAGES
PRONOMINAL PLURAL	no	yes	yes	yes
CLUSIVITY	no	no	no	no
NOMINAL PLURAL MARKING	no	optional	optional	(varied)
NUMERAL CLASSIFIERS	no	no	no	no
ASSOCIATIVE PLURAL	no	yes	yes	no
NOMINAL PLURAL MARKING	no	plural word	plural word	(as G or N input lg)
NOMINAL NUMBER AGREEMENT: MODIFIERS	no	no	no	(as G or N input lg)
NOMINAL NUMBER AGREEMENT: DETERMINERS	no	yes	yes	(as G or N input lg)
VERBAL NUMBER AGREEMENT	no	no	no	(as G or V input lg)
COORDINATING CONSTRUCTIONS REDUNDANCY	no	no	no	(as G or N input lg)
POLITENESS DISTINCTIONS	no	(optional)	(optional)	no
GENERICITY	base form	base form	base form	(data N/A)

number. Numerals tend to precede the noun, but it is variable if the noun remains in the bare form or is marked for number. Also similar to creoles is that while verbs are not marked for number the VP is number specified through the free subject pronoun.

In terms of discourse, pidgincreoles do not tend to mark politeness distinctions and tend to express genericity with the bare form of the noun.

Number is thus a much more relevant category in pidgincreoles than in pidgin languages, and in this respect pidgincreoles pattern much more closely to creole languages than to pidgins.

Mixed languages very commonly show a blend of the two (or few) input languages in their pronominal systems, and number tends to be indicated either through a free form or through affixation.

With respect to nominal number, mixed languages tend to reflect the system primarily of one of the input languages: G-L mixed languages tend to express nominal number in the same or in similar ways to the language which has provided most of the grammar, while N-V mixed languages tend to exhibit the same or similar nominal number marking strategies as the language which has provided the bulk of the NP. F-S mixed languages, on the other hand, tend to behave like the language which has provided most of the structure, but by means of markers taken over primarily from the language which has provided most of the forms. It seems somewhat more common for mixed languages to lack the associative plural than to have it.

Numeral classifiers seem marginal or not relevant for mixed languages, which is likely due to the fact that they are marginal or not relevant for the respective source languages.

Mixed languages vary considerably in terms of whether modifiers agree in number with the head noun. In some languages they do, including numerals, in others the head noun remains invariant. There is also considerable variation with respect to verbal morphology, and again the languages tend to behave the same as or similar to the input language which has provided the bulk of the grammar (for G-L languages), the bulk of the verb phrase (for N-V languages) or the bulk of the structure (for F-S languages).

Politeness distinctions are not generally found in mixed languages and there is little information available on how genericity is expressed.

Acknowledgements

Abbreviations

1, 2, 3	first, second third person / noun class number
ABS	absolutive
ACT	action marker
AGR	agreement
ANIM	animate
APiCS	Atlas of Pidgin and Creole Structures (https://apics-online.info/)
ART	article
ASP	aspect marker
APL	associative plural
AUX	auxiliary
COM	comitative
COMP	complementizer
COP	copula
DAT	dative
DEF	definite
DEM	demonstrative
DET	determiner
DIST	distal
DU	dual

ERG	ergative
EXIST	existential
F	feminine
FAM	familiar
FOC	focus marker
FUT	future
GEN	genitive
HAB	habitual
HON	honorific
INAN	inanimate
INDF	indefinite
INF	infinitive
IPFV	imperfetive
M	masculine
NC	noun class
NEG	negative
NMLZ	nominalizer
NPST	nonpast
ORD	ordinal
PAUC	paucal
PL	plural
PM	predicate marker
POL	polite
POSS	possessive
POT	potential
PREP	preposition
PROG	progressive
PROX	proximate
PRS	present
PST	past
REL	relative
SBJ	subject
SG	singular
TR	transitive

References

Aboh, Enoch O., Norval Smith & Tonjes Veenstra. 2013. Saramaccan. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds), *Survey of pidgin and creole languages. Vol. I: English-based and Dutch-based languages,* 27–38. Oxford: Oxford University Press.

Anonymous 1830. *Dictionnaire de la langue franque ou petit mauresque.* Marseilles: Typographie de Feissat Ainé et Demonchy.

Avram, Andrei. 2010. An outline of Romanian Pidgin Arabic. *Journal of Language Contact – VARIA* 3(1). 20–38.

Ayafor, Miriam. 2008. Cameroon Pidgin English (Kamtok): Morphology and syntax. In Rajend Mesthrie (ed.), *Varieties of English. Vol. 4: Africa, South and Southeast Asia*, 428–450. Berlin: Mouton de Gruyter.

Aye, Khin Khin. 2005. *Bazaar Malay: History, grammar and contact*. Singapore: National University of Singapore PhD dissertation.

Bakker, Peter. 2006. The Sri Lanka Sprachbund: The newcomers Portuguese and Malay. In Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic areas. Convergence in historical and typological perspective*, 135–159. Houndmills Basingstoke: Palgrave MacMillan.

Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 130–157. Oxford: Oxford University Press.

Bakker, Peter. 2020. Contact and mixed languages. In Raymond Hickey (ed.), *Handbook of language contact*, 2nd ed., 201–220. New York: Wiley.

Bhattacharjya, Dwijen. 2001. *The genesis and development of Nagamese. Its social history and linguistic structure*. New York: City University of New York PhD dissertation.

Bhattacharjya, Dwijen. 2007. Nagamese (restructured Assamese). In John Holm & Peter Patrick (eds.), *Comparative creole syntax. Parallel outlines of 18 creole grammars*, 237–254. London: Battlebridge.

Biagui, Noël Bernard & Nicolas Quint. 2013. Casamancese Creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/34, Last access 2018–12–18.)

Bollée, Annegret. 2013. Reunion Creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/54. Last access 2018–12–21.)

DeGraff, Michel. 2007. Kreyòl Ayisyen or Haitian Creole. In John Holm & Peter Patrick (eds.), *Comparative creole syntax. Parallel outlines of 18 creole grammars*, 101–126. London: Battlebridge.

Drechsel, E. J. 1997. *Mobilian Jargon. Linguistic and sociohistorical aspects of a Native American Pidgin*. Oxford: Clarendon.

Ehrhart, Sabine & Melanie Revis. 2013. Tayo. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages. Vol. II: Portuguese-based, Spanish-based and French-based languages,* 271–s82. Oxford: Oxford University Press.

Farquharson, Joseph T. 2013. Jamaican structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/8. Last access 2018–12–20.)

Fleury, Norman. 2007. La Pchit Sandrieuz, kii-aachimot Norman Fleury, Richmond, B.C., March 2004. Cinderella. Told by Norman Fleury, Richmond, B.C., March 2004. In Bakker, Peter (ed.) Aarhus/Saskatoon/Winnipeg: private publication, CD.

Foley, William A. 2013a. Yimas-Arafundi Pidgin. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Vol. 3: Contact Languages Based on Languages from Africa, Asia, Australia, and the Americas*, 105–113. Oxford: Oxford University Press.

Goddard, Ives. 1997. Pidgin Delaware. In Sarah G. Thomason (ed.), *Contact languages. A wider perspective,* 43–98. Amsterdam: John Benjamins.

Huber, Magnus. 2013. Ghanaian Pidgin English structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/16. Last access 2018–12–20.)

Intumbo, Incanha, Liliana Inverno & John Holm. 2013. Guinea-Bissau Kriyol structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/33. Last access 2018–12–21.)

Jabłońska, Antonina. 1957. *Język mieszany chińsko-rosyjski w Mandžurii (The Sino-Russian mixed language in Manchuria)*. Przegląd Orientalistyczny 21. 157–168.

de Jong, Josselin P. B. 1926. *Het huidige Negerhollandsch (teksten en woordenlijst)*. (Verhandelingen der Koninklijke Academie van Wetenschappen te Amsterdam, 26.) Amsterdam: Koninklijke Academie van Wetenschappen te Amsterdam.

Klein, Thomas B. 2013. Gullah. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages. Vol. I: English-based and Dutch-based languages,* 139–147. Oxford: Oxford University Press.

Kouwenberg, Silvia. 2013. Papiamentu structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/47. Last access 2018–12–20.)

Matthews, Stephen & Michelle Li 2013. Chinese Pidgin English. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages*. *Vol. I: English-based and Dutch-based languages,* 206–213. Oxford: Oxford University Press.

Maurer, Philippe. 1995. *L'angolar. Un créole afro-portugais parlé à São Tomé. Notes de grammaire, textes, vocabulaire*. (Kreolische Bibliothek, 16.) Hamburg: Buske.

Maurer, Philippe & the APiCS Consortium. 2013. Nominal plural marker and third-person-plural pronoun. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press. Feature 25. (Available at https://apics-online.info/parameters/25#2/30.4/10.2. Last access 2018–12–18.)

Meakins, Felicity. 2011. *Case marking in contact: The development and function of case morphology in Gurindji Kriol, an Australian mixed language*. Melbourne: University of Melbourne PhD dissertation.

Meakins, Felicity. 2013. Gurindji Kriol. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages*. *Vol. 3: Contact languages based on languages from Africa, Asia, Australia, and the Americas*, 131–139. Oxford: Oxford University Press.

Meakins, Felicity. 2013. Gurindji Kriol structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/72. Last access 2018–12–20.)

Mesthrie, Rajend & Clarissa Surek-Clark. 2013. Fanakalo. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Vol. 3: Contact Languages Based on Languages from Africa, Asia, Australia, and the Americas*, 34–41. Oxford: Oxford University Press.

Meyerhoff, Miriam. 2013. Bislama structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/23. Last accessed 2018–12–19.)

Michaelis, Susanne Maria & Marcel Rosalie. 2013. Seychelles Creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for

Evolutionary Anthropology. (Available at http://apics-online.info/contributions/56. Last access 2018–12–21.)

Migge, Bettina. 2013. Nengee structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/4. Last access 2018–12–20.)

Mous, Maarten. 2003. *The making of a mixed language: The case of Ma'a/Mbugu*. (Creole Language Library 26.) Amsterdam: John Benjamins

Muysken, Pieter. 1981. Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield & Albert Valdman (eds.), *Historicity and variation in creole studies*, 52–78. Ann Arbor: Karoma.

Muysken, Pieter. 2013. Media Lengua. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Vol. 3: Contact Languages Based on Languages from Africa, Asia, Australia, and the Americas*, 143–148. Oxford: Oxford University Press.

Paauw, Scott. 2008. *The contact Malay varieties of Eastern Indonesia: A typological comparison*. Buffalo, NY: University at Buffalo. Buffalo: University at Buffalo PhD dissertation.

Paauw, Scott. 2013. Ambon Malay. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Vol. 3: Contact Languages Based on Languages from Africa, Asia, Australia, and the Americas*, 94–104. Oxford: Oxford University Press.

Parkvall, Mikael. 2017. Pidgin languages. *Oxford Research Encyclopedia of Linguistics*. (Available online at https://doi.org/10.1093/acrefore/9780199384655.013.58. Last access 4 December 2020).

Perekhvalskaya, Elena V. 2008. *Russkie pidžiny (Русские пиджины)*. St. Petersburg: Aletejja.

Pfänder, Stefan. 2013. Guyanais structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/52. Last access 2018–12–21.)

Quindale, Francisco [. d. S. M.] 1867. *Diccionario gitano*. Madrid: Oficina Tipográfica del Hospicio.

Reinecke, John E. 1971. Tây Bòi. Notes on the Pidgin French of Vietnam. In Dell Hymes (ed.), *Pidginization and creolization of languages,* 47–56. Cambridge: Cambridge University Press.

Roberts, Sarah J. 2013. Pidgin Hawaiian structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/71. Last access 2018–12–21.)

Samarin, William J. 2013. Sango structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/59. Last access 2018–12–20.)

Schaengold, Charlotte. 2004. *Bilingual Navajo. Mixed codes, bilingualism, and language maintenance*. Columbus: The Ohio State University PhD dissertation.

Schultze-Berndt, Eva, Felicity Meakins & Denise Angelo. 2013. Kriol. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages. Volume I: English-based and Dutch-based languages,* 241–251. Oxford: Oxford University Press.

Slomanson, Peter. 2013. Sri Lankan Malay structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/66. Last access 2018–12–22.)

Smith, Geoff P. & Jeff Siegel. 2013. Tok Pisin structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/22. Last access 2018–12–22.)

Smith, Ian R. 2013. Sri Lanka Portuguese structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/41. Last access 2018–12–21.)

Smith, Ian R. 2013. Sri Lanka Portuguese. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages. Vol. II: Portuguese-based, Spanish-based and French-based languages,* 111–121. Oxford: Oxford University Press.

Solomon, Denis. 1993. *The speech of Trinidad. A reference grammar*. St. Augustine: UWI, The School of Continuing Studies.

Staudacher-Valliamée, Gillette. 2004. *Grammaire du créole réunionnais*. Paris: Sedes.

Steinkrüger, Patrick O. 2013. Zamboanga Chabacano structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), Atlas of pidgin and creole language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://apics-online.info/contributions/46. Last access 2018–12–21.)

Thomason, Sarah G. 1980. On interpreting 'The Indian Interpreter'. *Language in Society* 9. 167–93.

Tosco, Mauro & Owens, Jonathan. 1993. Turku. A descriptive and comparative study. *Sprache und Geschichte in Afrika* 14. 177–267.

Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: John Benjamins.

Velupillai, Viveka. 2015. *Pidgins, Creoles and Mixed Languages*. *An introduction*. Amsterdam: John Benjamins.

Verhaar, John W. M. 1995. *Toward a reference grammar of Tok Pisin: An experiment in corpus linguistics*. (Oceanic Linguistics Special Publications 26.) Honolulu: University of Hawai'i Press.

Volker, Craig A. 1982. *An introduction to Rabaul Creole German (Unserdeutsch)*. St Lucia: University of Queensland MA thesis.

Yakpo, Kofi. 2013. Pichi. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of pidgin and creole languages. Volume I: English-based and Dutch-based languages,* 194–205. Oxford: Oxford University Press.

Vadim Kimmelman, Svetlana Burkova, and Elizaveta Filimonova

22 Number in Russian Sign Language

Abstract: In this chapter, we describe how plural meaning is expressed in Russian Sign Language (RSL) as compared to other signed and spoken languages. The chapter is structured around the questionnaire developed for this volume. We demonstrate that RSL has means of expressing pronominal, nominal and verbal plurality. Moreover, RSL has a set of tools to distinguish various semantic subtypes of plurality. The main means used to express plurality are, in morphology, reduplication, incorporation, and modification of movement, as well as various lexical means such as quantifiers. Importantly, plurality is almost never obligatorily expressed. Some marker of plurality in a sentence is necessary to obtain the plural interpretation, but once one such marker is present, further plural marking on nouns and verbs is optional. Furthermore, no obligatory plural agreement in RSL is attested: multiple exponence of the plural meaning is possible and, in some cases, common, but never obligatory. While not much research has been done, plural markers in RSL seem to behave semantically and pragmatically in typologically expected ways. Finally, we observe that most morphological means of expressing plurality in RSL are highly iconic.

1 Overview

1.1 Plurality in sign languages

Number marking has been described for a few sign languages. Some sign language descriptions mention plurality albeit briefly (Sutton-Spence & Woll 1999; Valli & Lucas 2000; Zeshan 2000; Johnston & Schembri 2007; Bauer 2014) but a majority of sign languages, have not been described in this respect at all. At the same time, most studies devoted specifically to the category of number are focused on a single issue, such as nominal plurality (Miljan 2003), morphological marking of plurality (e.g. Pfau & Steinbach (2006)), reduplication and classifiers as markers of plurality (Nijhof & Zwitserlood 1999; Pfau & Steinbach 2006) and so on. Therefore, this topic clearly warrants further research.

Existing research shows that all known sign languages have some means of expressing the singular vs. non-singular opposition, while some sign languages show more differentiated oppositions, such as singular, dual, trial, quadruple, quintuple, paucal, and plural, as well as differentiate additive, collective, and distributive plurality. At the same time, at this stage it is not possible to talk about an obligatory grammatical category of plurality as applied to any word class (pronouns, nouns, verbs, classifier predicates). At least one sign language has been

claimed to have obligatory nominal number marking (Miljan (2003) for Estonian Sign Language, ESL), most sign languages appear to express plurality optionally (Pfau & Steinbach 2021), and for some sign languages the lack of regular marking of plurality has been argued for (e.g. Zeshan (2000) for Indo-Pakistani Sign Language, IPSL).

Several overviews of the expression of plurality across different sign languages exist (Pfau & Steinbach 2006; Pfau & Steinbach 2021). Therefore, the purpose of this chapter is not to provide another such overview. Instead, we focus on describing expressions of plurality in one sign language, namely Russian Sign Language (RSL), which has not been described before in such detail. We follow the questionnaire created for this volume to make sure that RSL can be compared to other languages and language families discussed here, although we do not have data for all questions. Wherever possible, we also place RSL in the typological perspective by comparing it to existing research on other sign languages. Note, however, that such comparison is often complicated by the inconsistent use of terminology and different methodologies applied across various studies.

The data reported in this chapter sometimes come from previous research (mostly by one or several of the authors of the chapter); in such cases, we refer to the source. In addition, we have conducted small-scale data collection in order to address some of the issues not previously investigated for RSL; whenever a source is not provided for an example, it means that it comes from data collection by the authors.

Before diving into the topic of number marking in RSL, we will describe some basic properties of sign languages, all of which are shared by RSL.

1.2 Basic properties of sign languages

Sign languages are natural languages existing in the visual modality (Sandler & Lillo-Martin 2006). In many respects, they are fundamentally similar to spoken languages. However, some features of sign languages are influenced by the visual modality, which is also the case for RSL. We only discuss some features relevant for the rest of the chapter.

Signs in sign languages have internal structure. Usually they are described as consisting of the handshape, orientation, location, and movement. Notably these components are combined simultaneously, although sequential structure might be present as well (if, for instance, there is a handshape, location, or orientation change in the sign). Hence morphological operations in sign languages are most commonly non-sequential: reduplication and stem modification are common, while affixation is extremely rare (Aronoff, Meir & Sandler 2005). This is also true for RSL.

Another level of simultaneity is added by the fact that sign languages including RSL use several articulators. The two main articulators are the two hands, which

are partially independent (Vermeerbergen, Leeson & Crasborn 2007). Some RSL signs are one-handed, in which case the second hand can be added for morphological purposes (such as expressing plurality), as we will show in the following sections.

In addition to the manual articulators, non-manual articulators (the body, the head, the mouth, the eyebrows, etc.) are used to express lexical and grammatical information (Pfau & Quer 2010). RSL uses these articulators to express negation, modality, to mark sentence types and information structure, among other things. As we show below, plurality can also be expressed non-manually.

Another important feature of sign languages is their use of space in front of the signer (the signing space). Referents (especially when they are absent) can be assigned to arbitrary locations in the signing space by adding a pointing sign or localizing a nominal or verbal sign in this location.[1] Thereafter these locations can be targeted to refer back to these referents by pointing signs used as pronouns and by verbal agreement. As we will show, the signing space plays a crucial role in expressing plurality in RSL.

Finally, sign languages have been shown to be highly iconic, both in the lexicon and in the grammar (Perniss, Thompson & Vigliocco 2010). As will become clear from discussion below, many of the means of expressing plurality in RSL are iconic, as there is a clear relation between the form of the sign/sentence and the meaning expressed.

2 Pronominal, nominal, and verbal number

2.1 Generalities

It is hardly possible to talk about a category of number in RSL (and apparently in a number of other sign languages) as an obligatory grammatical category. None of the values that belong to the semantic domain of number are included in RSL into any set of obligatory and mutually exclusive values. But some of them, e. g. those that indicate distributive plurality in nouns and verbs, can be regarded as quasigrammemes in terms of Mel'čuk (1993: 302–303), i.e. grammatical values expressed as regularly as grammemes without being obligatory. Plungian (2003: 133) assumes that, from a diachronic point of view, quasigrammemes present a stage immediately preceding a formation of a grammatical category. Sign languages are comparably young languages,[2] and the idea that they have precursors of full grammemes seems plausible.

---

1 Assignment of absent referents to arbitrary locations is not universal in sign languages, see e.g. De Vos & Pfau (2015).
2 See, for example, Brentari (2010) for an overview on history and transmission of sign languages.

Below we will discuss that morphological expression of plurality in RSL is almost always optional; most unmarked nouns and verbs can freely occur in plural contexts. However, in order for the semantic feature of plurality to be conveyed in a sentence, some marker of plurality is required. So, for instance, the sentence in (1) where neither the subject nor the verb are marked for plurality only has the singular interpretation.[3] In order to obtain the plural interpretation for the sentence, the signers can pluralize the noun, add a quantifier, or use a different verb that would encode number.

(1) BOY DANCE
'A boy dances'. NOT 'Boys dance'.

In this section, we discuss number marking in RSL pronouns, nouns, and verbs. Personal pronouns distinguish between the singular, dual, trial, quadruple and plural number expressed morphologically by means of reduplication and/or modification of a sign. Nouns are a heterogeneous class. There is a rather rich inventory of morphological and lexical means for marking number in nouns; their choice, as well as a set of values that can be marked in a certain noun, depends on phonological or morphological features of the sign, its lexical semantics, its origin, etc. Verbal number is also expressed by a variety of strategies partially dependent on the phonological and morphological features of the signs, as well as other factors.

2.2 Pronominal number

Unlike pronominal systems of spoken languages, pronominal systems of sign languages are fundamentally uniform (McBurney, 2002). Pronominal reference in these languages is highly associated with space, so personal pronouns are most often pointing signs directed towards to present referents or to locations in space associated with non-present referents (see McBurney 2002; Cormier 2012; Cormier et al. 2013).

The existing descriptions show that sign languages typically have rather rich systems of marking number in their pronouns. For example, American (ASL), Australian (Auslan), Danish (DSL), New Zealand (NZSL), and British Sign Language (BSL) have marking for the singular, dual, triple, quadruple, and plural number in their personal pronouns (McBurney 2002; Sutton-Spence & Woll 2007; Wallingford 2008). ASL and BSL are reported to have also quintuple forms (McBurney 2002: 336; Sutton-Spence & Woll 2007: 42–43).

Pronouns in RSL, similar to other sign languages, are pointing signs. The pointing sign in RSL has two variants: the first one is pointing with the extended index

---

3 See the end of the chapter for the Glossing conventions.

finger (the "1" handshape; palm downward or sideward, Figure 1a), and the second one is pointing with the tips of adjoining extended fingers of an open palm (the "B" handshape; palm upward, Figure 1b). The distribution of these two variants is not quite clear as of yet. In case of the 1st person both forms occur, although in the 1st singular the variant with the "1" handshape is more often preferred, while in the 1st plural – the variant with the "B" handshape. Non-first person pronouns also use both handshapes, but the variant with the "B" handshape apparently functions as a more polite form.

Marking number in RSL pronouns is very similar to that found in NZSL, Auslan and a number of other sign languages. There are forms for the singular, dual, triple, quadruple and plural derived morphologically from a pointing sign by means of reduplication or/and modification of a sign.

The reference to the 1st person[4] singular, as in many other sign languages, is most typically performed by pointing to the signer's chest, and the reference to the non-first singular – by pointing to a locus associated with the referent.

The form for the dual involves modification of handshape, palm orientation and movement: the hand with the palm located vertically moves forth and back between two loci associated with the referents. The hand can have two shapes: either "2" (extended index and middle fingers), or "L" (extended thumb and index finger),[5] both denoting 'two', see Figure 2. Note that the dual form is compatible with all possible persons: 'you and me', 'you and him', 'he and me', etc. In some contexts, the signers just use double pointing (with the "1" or "B" handshape) to the two loci associated with referents. However, it is not clear yet whether this should rather be considered as a phrase ('you and me', 'you and he/she', etc.), or as a single sign derived by reduplication accompanied by change in location at each repetition.

There are two forms of personal pronouns for the plural. The first one has an arc-shaped or sweeping movement (i.e. it involves modification of movement). This form is per se a collective pronoun and has the meaning 'we/you/they together'. The movement used in it iconically expresses the meaning of collectivity, when a group of referents is considered as a non-discrete homogeneous set. The same arc-shaped or sweeping movement is also a rather standard way of marking collective plurality in nouns, see Section 2.3; it is also used in the adverbial sign TOGETHER, verb sign GATHER, and the determiner ALL.

---

4 The analysis of person in sign language linguistics remains unsettled. Some earlier works assumed three-person systems of personal pronouns similar to those found in spoken languages (Friedman 1975; Padden 1988); McBurney (2002; 2004) and Liddell (2000) have argued for an analysis of pronominal systems that make no person distinctions; many researchers adopt the first / non-first model (Meier 1990; Lillo-Martin 2002). For further details see, e.g., Cormier et al. (2013). In this chapter, we follow the latter approach and distinguish between first and non-first person.
5 The choice between these two handshapes is unclear yet. It may either be due to idiolect variations or depend on the form of the adjacent signs.
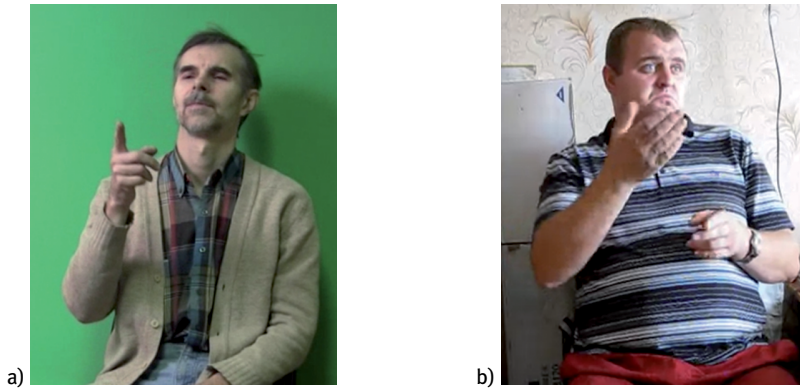
Fig. 1: Pointing signs in RSL a) with "1" handshape, b) with "B" handshape.
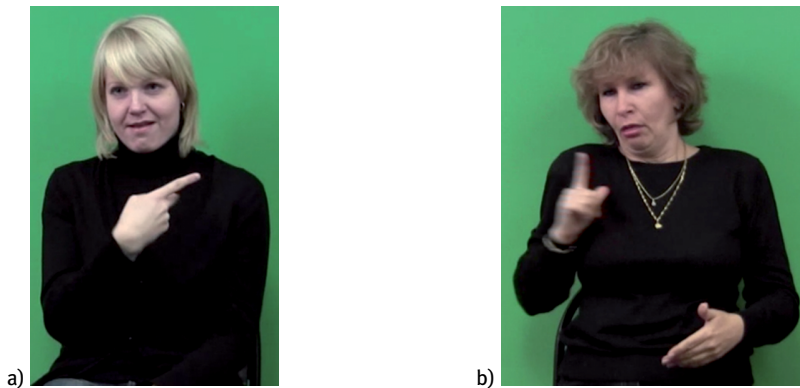
Fig. 2: Personal pronouns for the dual in RSL a) with "2" handshape, b) with "L" handshape.

The second form involves reduplication of a pointing sign accompanied by a change in location of the hand at each repetition. This form is per se a distributive pronoun and has a meaning 'each of us/you/them'. Reduplication accompanied by change in location is a standard way in RSL of marking distributive nominal and verbal plurality, when each referent (or situation) of a homogeneous set is considered distinct in space (or time), see also Sections 2.3 and 2.4.

The forms of personal pronouns for the triple and quadruple are similar in their movement to the form for the collective plural, but involve numeral incorporation: the handshape of a pointing sign ("1" or "B") is replaced with a numeral handshape "3" (Figure 3) or "4" respectively. We can assume that the forms for the triple and quadruple containing an arc-shaped or sweeping movement, similarly to that for the plural, express collectivity. However, they seem not to have distributive counterparts. In our observations, personal pronouns for the triple and quadruple in RSL,

Fig. 3: Personal pronoun for the triple in RSL.

as well as in a number of other sign languages, are optional. They are used when it is necessary to clarify that there are exactly three or four referents. Otherwise the form for the plural will be used.

We can see that the values that have overt and standard marking in the system of RSL personal pronouns (and, as may be expected, in a number of other sign languages), are the values of collective and distributive plurality.[6]

2.3 Nominal number

Nominal number marking has been described for many sign languages (see Pfau & Steinbach 2021 for an overview). The following strategies are commonly mentioned [1] morphological: reduplication, the use of the two hands, numeral incorporation, suppletion; and [2] lexical: the use of quantifiers or other lexical markers of plurality. Note that no systematic discussion of the different semantic types of nominal plurality (additive, collective, distributive, similative) in other sign languages exists.

In RSL, number in nouns is marked morphologically, by means of reduplication, modification of the sign, numeral incorporation, and suppletion, or lexically, by quantifiers and classifiers. The choice of exponence for a certain nominal sign mainly depends on the type of plurality being expressed, phonological and/or morphological structure of the sign, and its lexical semantics (Burkova 2015). The relevant phonological features include handedness (whether the sign is produced with one or two hands) and body-anchoredness (whether the sign is bound to some body part).

---

6 It is not clear yet what kind of plurality is expressed by the forms for the dual. However, a connecting movement between two locations used in these forms suggests that they are more likely to express a collective value 'we/you/they two together' than just a value of additive number.

The patterns of marking are quite distinctly distributed across different semantic domains of plurality: additive, collective, distributive, and similative plurality. Regularity and constraints on marking also correlate with different types of plurality.

2.3.1 Additive plurality

Additive plurality, that is, the meaning 'more than one X', in which "every referent of the plural form is also referent of the stem" (Daniel & Moravcsik 2013), is typically expressed morphologically, by means of simple or two-handed simultaneous reduplication. In the case of simple reduplication a sign is repeated without changes in its handshape, orientation, location or direction of movement. In case of two-handed simultaneous reduplication, an originally one-handed sign is produced with both hands moving simultaneously. The choice between the two types of reduplication is determined by phonological constraints. Simple reduplication is employed in case of one-handed body-anchored signs (e.g. DOCTOR in (2)) or two-handed signs in which the hands touch each other (e.g. FACT in (3)).

(2) KNOW PERSON TREAT DOCTOR+(r/s) THE.BEST / USELESS TREAT IMPOSSIBLE THE.SAME
'Even if the person had been treated by the best doctors, it wouldn't have helped anyway, [the disease is] incurable.'
(Burkova 2015: 176)

(3) IX$_1$ ALWAYS INTERNET INSIDE INTERESTING(r/dbl/sm) FACT+(r/s) TOPIC NAME INTERESTING FACT+(r/s) LOOK.AT SEARCH INTERESTING
'I am always online searching and looking for the most interesting news.'
(Burkova 2015: 176)

Two-handed reduplication is employed in case of one-handed non-body-anchored signs (e.g. PERSON and DOG in (4) or in case of two-handed non-body anchored signs in which non-dominant hand can be omitted without losing any important information (e.g. MOUSE in (5)).

(4) PAST+(r/s) PERSON(r/dbl/sm)   EVIL GOOD AND INCLUDING DOG(r/dbl/sm) / EVIL GOOD PAST+(r/s)
'There are evil and good people, the same with the dogs: there are evil and good ones.'
Source: http://rsl.nstu.ru/data/view/id/224/t/502160/d/510830[7]

---

[7] This and a number of other examples contain direct links to the source in the on-line corpus of RSL. Note that free registration is required to access the data.

(5) ONLY PERIMETER LITTLE DROP++++(r/s:m/arc) / DONE / MORE SMELL
    MORE MOUSE(r/dbl/sm) WON'T.BE / GO.AWAY OFF GOOD.BYE DONE
    'Just drop [the diesel fuel] along the perimeter. And that's it, say goodbye to
    mice, they will flee because of the smell.'
    Source: http://rsl.nstu.ru/data/view/id/12/t/59910/d/72120

However, both simple and two-handed reduplication are not specified for marking
nominal additive plurality; they can be used in other functions (see e.g. Section
2.4). Moreover, additive plurality does not necessarily have to be expressed by redu-
plication. Other means of expression can be also used.

One strategy for expressing pluralities with few members without resorting to
reduplication involves numeral incorporation. Sign languages use fingers as the ba-
sis for building numerals – this is a natural effect of the fact that hands are the
primary articulators. In addition, many sign languages use the handshapes repre-
senting numerals in combination with various movements and locations in order to
express other concepts related to quantity, see for instance Zeshan et al. (2013). For
example, the RSL sign HOUR contains a circular movement with a fist handshape,
but the sign can also be produced with a handshape with the thumb and index
outstretched to convey 'two ours', or with the thumb, index and middle fingers for
'three hours', etc.

This is usually called numeral incorporation, as it is often possible to find a
lexical sign which in isolation means some concept (for instance, *hour* or *week*), but
which can incorporate the numerical handshape in order to express quantity (Meir
2012). However, this way of expression additive number values is restricted to cer-
tain lexical domains (their list appears to be almost universal in different sign lan-
guages), and also has constraints related to the anatomical structure of the hand.

Thus, numeral incorporation is employed in RSL in the following domains: time
(minutes, seconds, hours (both duration and the time of the day), days (+ over *n*
days), weeks, months, years (+ *n* years back)), nominal classifiers (pieces, persons,
times), money (roubles), and also kilograms and places (that is, medals in sports).
Moreover, expressions like "with the *n* of them" can be analysed as numeral incor-
poration as well.

Incorporation is partially determined by the phonological form of the RSL nu-
merals. Numerals from ONE to FIVE are one-handed (and thus can be easily incorpo-
rated), numerals from SIX to TEN are two-handed, and all the numerals above TEN
contain lexical movement (and thus are almost impossible to combine with another
sign). The general rule is therefore that numerals from ONE to FIVE are incorporated,
while no other numerals are. This is the case for instance with the sign DAY: there
are signs ONEDAY, TWODAY, THREE$DAY, FOUR$DAY, and FIVE$DAY, and if the
quantity of days is larger, then a combination of a numeral and a lexical sign DAY
is used. However, there are some exceptions to this general pattern. For instance,
the noun HOUR (referring to duration) can incorporate numerals up to TEN, and

there are also archaic forms with ELEVEN and TWELVE incorporated, and the noun MONTH can incorporate numerals up to TEN.

Some nouns have a suppletive plural form, that is, a sign that expresses the plural meaning but that is not phonologically related to the sign expressing the singular meaning, e.g. CHILD – CHILDREN, HAIR – HAIR.PL, TOOTH – TEETH.

2.3.2 Collective plurality

Collective plurality, in which all referents of a plural form are conceptualized as a non-discrete homogeneous unit, is regularly expressed by smooth arc-shaped or sweeping movement, typically produced horizontally. Non-body-anchored signs are produced with the arc encompassing movement by themselves, e.g. TREE(m/arc) 'forest', BUSH(m/arc) 'shrubbery', THREE(m/arc) 'all three', FIVE(m/arc) 'all five', etc. In case of body-anchored signs which do not allow change in location, a periphrastic construction should be used instead: a lexical sign produced in a citation form is followed by the pointing sign IX produced with the arc movement, e.g. PROFESSOR IX(m/arc) 'professorate', AGENT IX(m/arc) 'agents', see also (6), where the body-anchored sign MAN is not modified morphologically, but followed by the pointing sign with arc movement.

(6) C-Ш-A[8]   MAN IX(m/arc)   ARMY SERVE MUST FREE
     'The men in the USA are not obliged to serve in the army'.
     (Burkova 2015: 179)

2.3.3 Distributive plurality

Distributive plurality indicates a set of spatially distributed objects. In case of non-body-anchored signs, it is typically marked by simple or two-handed successive[9] reduplication both accompanied by change in location at each repetition along arc-shaped path. This combination of reduplication with change in location at each repetition appears to be specialized for distributivity. Apart from distributive plurality in nouns, it is also used to express the distributive aspect in verbs, see Section 2.4. The choice between the two structural types of reduplication (simple vs. two-handed) is determined by the sign's phonological structure. Simple reduplication is mainly employed in two-handed signs (as ROOM in (7)), whereas two-handed successive reduplication is mainly used in case of one-handed signs (as PUDDLE in (8)).

---

8 A fingerspelled Russian abbreviation *США* 'USA'.
9 In case of two-handed successive reduplication, an originally one-handed sign is produced with both hands moving successively.

(7) DOCTOR ROOM SMALL ROOM++(r/s:m/arc)
    'There is a small hospital with several wards.'
    Source: http://rsl.nstu.ru/data/view/id/220/t/41410/d/43030

(8) STREET ALL L-U-ZH-A PUDDLE+++(r/dbl/alt:m/arc)
    'There are puddles in the street.'
    (Burkova 2015: 181)

In body-anchored signs that do not allow change in location, the two-handed successive reduplication is used, as in (9). The same structural type of reduplication is also employed in case of "large" non-body-anchored signs. For example, the non-body-anchored sign MOUNTAIN in (10) has a rather large path of movement taking the most part of signing space, so to change the location of this sign is not easy.

(9) LOOK FIELD FLOWER++(r/dbl/alt)
    'He saw a field, all in flowers.'
    Source: http://rsl.nstu.ru/data/view/id/39/t/28740/d/31510

(10) MOUNTAIN+(r/dbl/alt) PAST IX
    'Have you ever been to the mountains?'
    (Burkova 2015: 181)

It is important to note that plurality marking, and especially distributive plurality marking, is often combined with, or is part of the expression of locative information about the objects referred to. For instance, a signer might describe chairs standing in a circle, or in two lines opposite each other, or on top of each other, using reduplication and spatial modification of the sign for CHAIR in an iconic way. Various analyses of such marking and the interaction between plurality and iconic spatial information have been proposed in the literature; some researchers arguing that even simple reduplication expressing additive plurality has an iconic component (Schlenker & Lamberton 2019). Clearly more cross-linguistic research of this issue is necessary.

2.3.4 Similative plurality

Some languages express associative plurality, that is, the meaning 'X and X's associates', in contrast to other types of nominal plurality, indicates a heterogeneous set of objects, see, for example, Daniel & Moravcsik (2013) for further details. RSL does not have a strategy of expressing associative plurality, but it does have a strategy expressing a related heterogenous plurality type, namely similative plurality: 'X and things similar to X'. This type of plurality is expressed in RSL periphrastically: a

lexical sign or several lexical signs that denote one or several heterogeneous referents of a set are followed by the sign DIFFERENT, e.g. BUS TROLLEYBUS DIFFERENT 'public transport facilities', DEAF DIFFERENT 'deaf and hard-of-hearing people', TOMATO CUCUMBER DIFFERENT 'vegetables', CHAIR BED DIFFERENT 'furniture', PLATE DIFFERENT 'tableware'.

It is important to note that this way of expressing similative plurality in RSL is not productive: only a few concepts are expressed this way, and, by Burkova's (2015) observations, it is more typical of older signers, whereas younger signers prefer to express the concepts like 'furniture', 'vegetables', etc. by fingerspelling the corresponding Russian words. There is also variation in acceptability of different variants of the same concept: for 'vegetables' some signers accept TOMATO CUCUMBER DIFFERENT, some only TOMATO DIFFERENT, and a few also accept CUCUMBER DIFFERENT.

2.3.5 Summary

As one can see, morphological means used to express different kinds of nominal plurality in RSL are highly iconic. It seems rather natural to associate reduplication of a sign with a discrete set of objects, whereas smooth arc movement seems to be naturally associated with non-discrete set of objects, as well as change in location at each repetition of a reduplicated sign seems to be naturally associated with the distributive localization of the objects of a set. For further discussion of the role of iconicity in plural marking see for instance Schlenker & Lamberton (2019).

2.4 Verbal number

In many sign languages, verbs can be modified to express plurality of different types, as first described for ASL by Klima & Bellugi (1979). Almost universally, the following possibilities exist: (1) simple reduplication is used to express plurality of events (e.g. iterative), as for instance in ASL and French Sign Language (Kuhn & Aristodemo 2017). (2) Arc-shaped movement is used to express plurality of arguments, as in Spanish Sign Language (Costello 2016). (3) Arc-shaped movement combined with reduplication is used to express distribution over arguments (Klima & Bellugi 1979; Tsay & Myers 2009; Kimmelman 2015; Filimonova 2012; Filimonova 2016). (4) Some information about the plurality of arguments can be encoded in classifier predicates, that is, predicates of movement in which the handshape refers to the class of an argument, while location and movement components are used iconically to encode location and/or motion of this argument (Zwitserlood 2012).

Partially based on the similarities between different sign languages in this domain, Wilbur (2010) has argued for the Event Visibility Hypothesis stating that certain properties of events are iconically encoded in sign languages. Below we demon-

strate that all of this also holds for RSL. However, our research shows that the patterns of expressing verbal plurality are quite intricate and complicated, and one should not expect to find the exact same patterns across different sign languages even if iconic mechanisms are involved.

2.4.1 Number in lexical verbs

Number marking on verbal signs in RSL is not obligatory. Example (11) below shows that while the subject is plural (more specifically, dual), no marking on the verb is present.

(11) **COME** FIRST.AID.STATION **COME**
'[Me and my friend] came to first-aid station'.

However, in RSL, special verb forms exist to express distributive and collective plurality (of subject/object), iterative (plurality of time periods in which an action is performed) and the reciprocal meaning (plurality of participants who are subjects and objects of the same situation).

2.4.1.1 Distributive and collective plurality›

As discussed in the previous section for nominal signs, distributive plurality in verbal signs in RSL is expressed by different types of reduplication and modification of the sign. The choice of reduplication type and modification first of all depends on whether the subject or object of the situation is plural (Filimonova 2016: 238–241).

Plurality of the subject can be conveyed by simultaneous or successive reduplication. Simultaneous reduplication is used when participants of the situation perform action simultaneously (12) and successive reduplication is used when they do it one by one or in different periods of time (13).

(12) VISIT$_1$ (r/dbl/sm)
'Guests came to us [at the same time]'.
(Filimonova 2016: 249)

(13) IX$_1$ BIRTHDAY ANNOUNCE GUESTS EIGHTEEN NULL NULL / IX TRAFFIC.JAM
IX BUSY IX DO.SOMETHING VISIT$_1$ (r/dbl/alt)
'I invited friends to come to celebrate my birthday at 18.00, but one of them got stuck in traffic jam, the other one is busy, the third has got something else, so they came one by one'.
(Filimonova 2016: 249)

Furthermore, arc movement can be used to convey non-distributive (collective) plurality of the subject, as in (14):

(14)   WHITE(m/ar)
       'Everything is white.'
       (Filimonova 2016: 251)

Plurality of the direct object is conveyed by arc movement or reduplication combined with change in location with each repetition. The latter is much more common in RSL. We suppose that the difference between these two marking strategies depends on whether the participants of situation are individuated (15) or seen as a collective participant (16).

(15)   PERSON INDX PAST LONG.AGO PERSON:PL KILL++(m/arc) /
       PERSON:PL GUILTY NOT IX KILL++ (r/s:m/arc) PAST
       'This person killed many people long time ago. These people were innocent and he killed them one by one'.

(16)   MULTIPLE.SUBJECT.MOVE KILL(m/arc) // KILL(m/arc) SAVE COULD /
       CAPTAIN GIVE.ORDER MAKE.MISTAKE
       'Many people were killed. They could have been saved but captain gave the wrong order'.

These modifications of the sign have some phonological restrictions. Body-anchored signs cannot undergo simultaneous, successive reduplication, arc movement or change in location because they are 'attached' to a certain place of signer's body. Double reduplication is possible only for one-handed signs or two-handed signs that can be reduced to one-handed ones. Successive reduplication is impossible for the signs that already have alternating movement in their structure.

As we mentioned above, other sign languages use the same means to express distributive and collective plurality: two-handed successive reduplication[10] (ASL), arc movement (ASL, Taiwan Sign Language: TSL), and a combination of reduplication with a change in location with each repetition (ASL, TSL, IPSL, ESL). However, some cross-linguistic variation can be observed.

For instance, a combination of reduplication with a change in location in ASL is used to express both subject and object plurality (Klima & Bellugi 1979), while in TSL and RSL it conveys only object plurality (Tsay & Myers 2009: 90). Furthermore, in ASL and BSL, arc movement is used to indicate collective meaning, and two-handed successive reduplication is used to express distributive meaning (Klima &

---

10 Also known in the literature as alternating reduplication.

Bellugi 1979; Sutton-Spence & Woll 1999). In TSL, the two forms are differentiated based on the period of time in which the actions are performed: if they occur simultaneously, arc movement is used; otherwise two-handed successive reduplication is used (Tsay & Myers 2009: 90).

2.4.1.2 Iterative

In all sign languages where the iterative has been studied so far (ASL, Swedish Sign Language, IPSL, Israeli Sign Language: ISL, Auslan) researchers have found that the basic means of expressing iterative meaning is simple reduplication without change of location (Klima & Bellugi 1979; Bergman & Dahl 1994; Zeshan 2000; Maroney 2004; Reagan 2007; Johnston & Schembri 2007; Meir & Sandler 2008). Also, for some sign languages it has been noted that some additional means can specify iterative meaning: movement modifications, speed and number of repetitions, and non-manual markers.

Iterative meaning in RSL is expressed by simple reduplication of the verb or the reduplicated form of the sign PAST. Reduplication is used only on non-stative predicates. Our data show that most of these reduplicated forms have one repetition. We suppose that two or more repetitions are used to emphasize plurality of situations (Filimonova 2016: 229–231).

(17) MY CLASSMATE+(r/s) OFTEN CINEMA **COME+**(r/s) TOGETHER
'Me and my classmates often go to the cinema'.
(Filimonova 2016: 230).

(18) **INJECT++**(r/s) DAY MANY
'[I was in hospital three months]. I got injected many times'.
(Filimonova 2016: 231)

Reduplicated sign PAST is a special iterative marker that has no limitation on the situation type. The initial form of the sign can serve as existential verb or auxiliary verb. The reduplicated form has the meaning 'it happens' or 'it happened from time to time'.

(19) **PAST+**(>r/s) **HEAD.HURT PAST+**(r/s)
'My head hurts sometimes'.
(Filimonova 2016: 232)

2.4.1.3 Reciprocals

Reciprocal meaning in RSL is expressed by special modifications of the verbal sign, which can differ depending on the number of participants. If there are only two participants in a reciprocal situation, a specific type of reduplication is used. (20)

illustrates a simultaneous two-handed mirror reduplication: both hands are specified for the same handshape and are situated in the signing space symmetrically and mirroring each other.

(20) FIRST TIME SEE-REC(r/dbl/sm:mirr) PLACE CONTEST DANCE
      'We saw each other for the first time at dance contest'.

In (21) backward reduplication is used. Both hands move back and forth with each repetition.

(21) WORK CAN TOGETHER HELP-REC(r/s:inv)
      'We can help each other in work'.

If there are more than two participants in the situation, we can see another modification of the sign: circular movement (22).

(22) IX$_{1PL}$ INDX(r/dbl/alt:m/crcl) INDX(r/dbl/alt:m/crcl) CAN HELP-REC(m/crcl)
      EACH.OTHER INDX(r/dbl/alt:m/crcl) WORK
      'All of us can help each other at work'.

The same differentiation of reciprocal forms with respect to number of participants has been found in German Sign Language (DGS). Pfau and Steinbach state that for reciprocal situations with more than two participants "randomized" (not directed towards specific locations in the signing space) movement is used (Pfau & Steinbach 2003).

2.4.2 Classifier predicates

Another way to express plurality on the predicate as described for many sign languages is to use a classifier predicate (Zwitserlood 2012). These are predicates of movement in which the handshape refers to the class of an argument, while location and movement components are used iconically to encode location and/or motion of this argument. For example, the sign with the "1" handshape (index finger) can refer to a person moving, while using the same sign with the "5bent" handshape (all fingers slightly bent) can refer to a round object, e.g. a ball, moving.

In addition to conveying information about the type of subject/object, the handshape can also convey information about the number of entities involved. In RSL, there are at least two ways of expressing plurality in classifier predicates. First, for human beings, as well as long-shaped objects, the "1"-handshape classifier can be used; to express multiple people or objects moving or located somewhere, the hand-

shape can change to include the corresponding number of fingers: 2–5 (similar to what happens in numeral incorporation), as illustrated by (23):

(23) THREE STREET CL:PERSON.THREE-MOVE
  'The three people are walking down the street.'
  Source: http://rsl.nstu.ru/data/view/id/31/t/40600/d/42810

Secondly, RSL has a classifier handshape reserved specifically for plural entities, namely the classifier MASS (made with two hands palms downwards, handshape "5"). It indicates a plural argument of many kinds: groups of people (25)–(27), flight of birds (24), as well as quantities of materials like flour or sugar (28).

  This sign can be modified in various ways by movement and orientation. In (26) circular movement is used to describe a situation where a group of people are walking around the building. In (27) the hands move from different directions to one place to indicate the movement of participants, and in (28) the hands move from one place to different directions.

(24) C-R-O-W MANY CL:MASS-MOVE BEAK EAT TASTY BEAK
  'Many crows flew [to the tree] and beaked [berries]'.
  http://rsl.nstu.ru/data/view/id/47/t/11190/d/11720

(25) IX$_1$ WHEN IX$_1$ TIE.SHOELACES IX$_1$ SIT TIE.SHOELACES HAPPEN
  GUN.SHOOT CL:MASS-MOVE IX$_1$ CAN BE.LATE RUN
  'When I was tying my shoelaces, [someone] shot starter gun and everyone ran, I could be late so I started to run too'.
  http://rsl.nstu.ru/data/view/id/231/t/493800/d/494280

(26) CL:MASS-MOVE.AROUND
  'Crowd walked around the building'.
  http://rsl.nstu.ru/data/view/id/218/t/218930/d/221840

(27) DIFFERENT MANY ABROAD CL:MASS-MOVE.TO.ONE.PLACE MOSCOW
  'A lot of people came to Moscow from different countries'.
  http://rsl.nstu.ru/data/view/id/198/t/425190/d/425440

(28) MOTHER F-L-O-U-R TAKE CL:MASS-MOVE.IN.DIFFERENT.DIRECTIONS
  'Mother took a bag of flour and flour scattered'.
  (Filimonova 2016: 248)

2.4.3 Number vs. aspect

The domain of aspect in RSL is very rich, as many subtle aspectual meanings can be expressed by morphological and lexical means. It is important to note that, ap-

parently, there is no perfective/imperfective opposition in RSL and perfective/imperfective interpretation is generally determined by the class of predicate: states are always seen as imperfective, achievements are always perfective, and processes and accomplishments are interpreted in the context (Filimonova 2016: 270). There are special markers to express habitual meaning and duration. There is no regular marker for durative situations although a few verbs referring to achievement or accomplishments can derive signs referring to processes by deceleration of movement.

Simple reduplication is used to express the iterative, habitual and durative, but in case of latter two it is combined with other lexical markers, such as the sign ALWAYS. Furthermore, there is a special marker for expressing longer than expected duration of a situation – non-manual reduplication (head shaking or rocking of the body). It can be combined either with simple reduplication (29) or with hand perseveration (30) (Filimonova 2016: 157–160). In the latter case a hand performing the sign stays fixed and does not move while non-manual reduplication is used.

(29)   ____bodyrocking
  BOY CRY+
  'Boy was crying for a long time'.

(30)   ____headshaking
  BOY THINK /   ASK
  'Boy was thinking (lit.: thinking, thinking) for a long time and then asked [the girl]'.

2.4.4 Interim conclusions

As discussed in sections 2.2–2.3, nominal and pronominal number, like verbal number, make use of different types of reduplication and modifications of a sign. They are very similar in expressing distributive meaning. Double successive reduplication is used to express distribution of some entities. Arc movement is used in to express collective plurality in both nouns and verbs; a combination of change in location and reduplication is used to express distributive plurality. Importantly, in all sign classes the phonological shape of the signs partially determines what types of reduplication can occur. It may seem at first glance that verbal and nominal number differ in the use of simple reduplication: while reduplicated verbs have iterative meaning, reduplicated nouns have the meaning of additive plurality. One could however argue that the iterative meaning in the verbal domain is the direct semantic parallel of additive plurality in the nominal domain.

3 Agreement and the syntax of number

It is probably safe to claim that no sign language demonstrates true obligatory agreement in number either within the NP domain, or between the verb and its arguments (see also Pfau & Steinbach 2021). RSL is no exception. For some sign languages it has even been claimed that certain configurations with multiple exponence of plural (e.g. reduplication used on the subject and plural marking on the verb in DGS (Pfau & Steinbach 2006)) are prohibited. RSL does not appear to prohibit multiple exponence of the plural number, but it also clearly does not systematically demand such multiple exponence.

We start with agreement within the NP, specifically, with plural marking on adjectives. For some sign languages, it has been claimed that adjectives cannot agree in number with the head noun; moreover, they cannot show plural marking at all (DGS, Pfau & Steinbach (2006)). In other sign languages, adjectives can express plural, also when the head noun is marked with plural, but not obligatorily (e.g. ESL, Miljan (2003), VGT, Heyerick et al. (2011)). In RSL adjectives are sometimes reduplicated (31), but, according to Burkova & Filimonova (2014), reduplication of adjectives is used not to mark plurality, but intensive instantiation of the characteristics as is clear from context in corpus examples, and from discussion with native signers. Thus, this cannot be analyzed as agreement.

(31) INTERESTING + FACT-PL
      'very interesting facts'
      (Burkova & Filimonova 2014: 230)

Further, the plural meaning in most sign languages can be expressed by quantifiers, including numerals, and the head noun would then be either optionally marked with morphological plural marking (as in ISL (Stavans 1996), Hausa Sign Language (Schmaling 2000)), or obligatorily unmarked for plural (as in DGS and ASL (Pfau & Steinbach 2006)). In RSL, plural marking on the noun in the presence of quantifiers is possible, albeit not obligatory (Kimmelman 2017), as shown by example (32).

(32) MANY QUESTION / MANY QUESTION-PL
      'many questions'
      (Kimmelman 2017: 824)

Kimmelman (2017) has argued that for some nouns, such as TOOTH and RIB, plural marking to express plural meaning is not optional (probably because it is expressed by a suppletive strategy). Such nouns must be marked morphologically as plural also when combined with quantifiers (33).

(33) *SOME TOOTH / SOME TEETH
      'some teeth'
      (Kimmelman 2017: 824)

However, if the noun is left-dislocated and clearly marked as a topic, the restrictions do not apply (34)–(35). Compare this to English "As for children, I only have one". The same pattern can be observed with regular nouns marked for plural: while in noun phrases, they cannot combine with the numeral ONE (36), if the noun is topicalized, no conflict occurs (37).

(34) <u>top</u>
RIB   SOME
'some ribs'
(Kimmelman 2017: 824)

(35) <u>   top</u>
CHILDREN   ONE
'one child'
(Kimmelman 2017: 824)

(36) *ROOM EXIST ONE CHAIR-**PL**

(37) <u>   top</u>
CHAIR-PL   ROOM EXIST ONE
'There is one chair in the room.'

Moving beyond the NP, some sign languages, such as VGT, NGT, and Turkish Sign Language (Heyerick et al. 2011; Nijhof & Zwitserlood 1999; Zwitserlood, Perniss & Özyürek 2012) have been argued to allow multiple exponence of plurality in the clause, such as for instance marking on the subject and on the verb, while others, such as DGS (Pfau & Steinbach 2006) seem to prohibit it. RSL again clearly allows marking plurality on arguments (either morphologically or by a quantifier) and the verb simultaneously, as demonstrated in (38), but this is clearly optional, as demonstrated in (39). Reduplication expressing reciprocal meaning, can also optionally occur on both the verb and the argument (40).

(38) EARLIER SICK TUBERCULOSIS PERSON-PL DIE-DISTR MANY
'Earlier, many people died of tuberculosis.'
(Burkova & Filimonova 2014: 236)

(39) LATER BEGIN PERSON COME-DISTR LIVE
'Later people started moving in.'
(Burkova & Filimonova 2014: 236)

(40) IX$_1$ PROMISE-REC IX-REC MORE SAME LOSE FUT.NEG
'We promised each other not to give up.'
(Burkova & Filimonova 2014: 242)

Multiple exponence of plurality is especially visible in distributive plural contexts. Kimmelman (2017) discusses example (41) where the subject NP is marked as plural by a lexical quantifier EVERY, as well as by a plural pointing sign IX-PL, and the verb is also marked by the morphological distributive marker. Morphological distributive marking can also occur on numerals (42), and even on the quantifier EVERY itself. However, multiple exponence of distributive plural marking is optional (42).

(41) EVERY BOY IX-PL DISTR-GIVE.PRESENT$_1$
      'Every boy gave me a present.'
      (Kimmelman 2017: 818)

(42) MAN BUY BEER ONE-DISTR
      'Every man bought a beer.'
      (Kimmelman 2017: 820)

Finally, a very interesting case of multiple exponence concerns non-manual marking of plurality (recall examples (29)–(30) above). Burkova & Filimonova (2014) demonstrate that in some cases non-manual markers of plurality (head or body movements) co-occur with manual morphological marking of plurality (reduplication) of the sign (43). However, non-manual marker of plurality can also occur on its own (44).

(43)   bht+
      HANDICAPPED WHEELCHAIR BAG-PL   HOLD-PL
      'I am in the wheelchair, holding the bags.'
      (Burkova & Filimonova 2014: 238)

(44) bht+
      WAIT NEG.EXIST
      'They waited and waited, but he was not there.'
      (Burkova & Filimonova 2014: 238)

To sum up, while RSL (and some other sign languages) allows multiple exponence of the plural meaning both inside the NP, and at the clause-level, this is always optional, and thus would not qualify as prototypical agreement. Interestingly, some other sign languages, most notably DGS, have been claimed to disallow multiple exponence of plurality.

4 Semantics and discourse

Very little research has been done on the semantics and pragmatics of plural marking in sign languages (see Kuhn & Aristodemo (2017) for some discussion of the

semantics of verbal plural marking, and some references below). For RSL, almost no research has been done. However, some aspects which have been studied are discussed in this section.

While many spoken languages use plural forms of pronouns as honorifics, as far as we know, no sign language does so. For NGT and ASL it has been found that a different handshape in pointing sign is used in polite contexts, namely the flat hand handshape instead of the index finger (Baker & van den Bogaerde 2012; Neidle & Nash 2012). The same is true for RSL: normally, pointing signs use the index finger, but the flat hand is used for politeness, as discussed in Section 2, see also Figure 1. Note that this handshape is not used to express the plural meaning as such, and also that the use of this handshape is optional.

Another common function of plural marking in spoken languages is to express generic reference. This is also attested in some sign languages (see Barberà and Cabredo Hofherr 2018 for an overview of impersonal reference in sign languages). Kimmelman (2018) found that plural pronouns can also be used in universal generic contexts in RSL, as in (45). However, the most common way to express generic (and, more broadly, impersonal) reference in RSL is to omit the subject (46).

(45)   <u>br</u>
    CHINA,   IX-PL LAST WEEK NEW YEAR IX-PL CELEBRATE
    'In China, they celebrated New Year last week.'
    (Kimmelman 2018: 220)

(46) GERMANY, EIGHT HOUR START WORK
    'In Germany, they start working at 8 o'clock.'

Plural marking on nouns in generic contexts is also possible, as in example (47) where the generic subject PERSON is marked both morphologically, by a lexical plural marker MASS, and by the universal quantifier ALL.

(47) PITY ALL SEE PERSON-PL MASS OLD
    'It's a pity to see that all people become old.'
    Source: http://rsl.nstu.ru/data/view/id/231/t/1602730/d/1606320

For nouns with suppletive plural, the plural form is required to obtain the relevant interpretation in generic contexts: compare (48) and (49).

(48) IF MANY CANDY EAT, TOOTH FUT HURT
    'If you eat too much candy, your tooth will hurt.'

(49) IF MANY CANDY EAT, TEETH FUT HURT
    'If you eat too much candy, your teeth will hurt.'

In some languages, such as Russian, plural marking on verbs can be used in generic contexts. While some generic contexts in RSL also use plural marking on the verb, Kimmelman (2018) argued that this marking is not related to genericity; it simply marks the fact that the event takes place multiple times, as in (50).

(50) NEIGHBORHOOD BIKE STEAL-ITER OFTEN
     'They often steal bikes in the neighborhood.'
     (Kimmelman, 2018: 213)

In discourse, when plurality is marked (by a lexically plural noun, or morphologically, or by a pointing sign or another lexical marker of plurality), the plural noun phrase can still admit singular reference in questions, as in (51)–(52).

(51) A: CHILDREN EXIST?   B: YES, ONE
     'A: Do you have children? B: Yes, I have one.'

(52) A: ROOM PERSON-PL EXIST?   B: YES, ONE PERSON
     'A: Are there people in the room? B: Yes, there is one person.'

Finally, with respect to the question of whether plural marking is used for reference tracking in discourse very little can be said. It is clear that, in RSL, plural pronouns can be used to disambiguate discourse referents, as in example (53). Another device clearly used to track referents are classifier handshapes, and sometimes these handshapes also include number marking. For instance, in (23) repeated here as (54), the referents (three men) were introduced in the previous sentence, and they are referred to again by a numeral and the classifier handshape, so number marking is (indirectly, and in combination with other markers) used to track referents.

(53) PERSON-PL KIND SAME IX-PL ATTENTION
     'If people are kind, it (the dog) is kind (lit. attentive) towards them.'
     Source: http://rsl.nstu.ru/data/view/id/224/t/522220/d/524810

(54) THREE STREET CL:PERSON.THREE-MOVE
     'The three people are walking down the street.'
     Source: http://rsl.nstu.ru/data/view/id/31/t/40600/d/42810

We are not aware of any in-depth discussion of discourse-related use of number in other sign languages.

5 Conclusions

In this chapter, we described how plural meaning is expressed in RSL – and, by comparison, how it is expressed in some other sign languages. We can make the following main conclusions.

First, RSL has means of expressing pronominal, nominal and verbal plurality. Moreover, RSL has a set of tools to distinguish various semantic subtypes of plurality, namely additive, collective, distributive, and similative plurality. The main means used to express plurality are, in morphology, reduplication, incorporation, modification of movement, as well as various lexical means such as quantifiers.

Second, plurality is almost never obligatorily expressed. Some marker of plurality in a sentence is necessary to obtain the plural interpretation, but once one such marker is present (e.g. a quantifier), further plural marking on nouns and verbs is optional. The exception to this pattern are signs with suppletive plural forms which have to be plural in plural contexts. However, even those signs can be realized as singular when they are topicalized and thus are not a part of the noun phrase with a plural quantifier.

Third, no obligatory plural agreement in RSL (or most if not all other sign languages) is attested. Multiple exponence of the plural meaning is possible and, in some cases, common, but never obligatory.

Fourth, while not much research has been done, plural markers in RSL seem to behave semantically and pragmatically in typologically expected ways.

Finally, we have observed that most morphological means of expressing plurality in RSL are highly iconic, which might be the reason why RSL strongly resembles other sign languages in this domain. However, we argue that the specific details of plural marking show clear cross-linguistic variation, and thus more research on different sign languages is necessary.

Glossing conventions

Signs are glossed in small caps; fingerspelled sequences are separated by dashes. The . is used when a single sign is glossed with more than one word; ^ connects components of compounds; $ marks numeral incorporation. IX stands for index (pointing signs). CL stands for classifier, followed by : and the meaning of the classifier handshape. / marks a prosodic boundary. Personal marking on pronouns and verbs is expressed by subscripts. In section 2, reduplication types are glossed in detail. The number of +'s reflects the number of repetitions.

Abbreviations

Abbreviations in section 2

r	reduplication
m	movement modification; these symbols are followed by a / and further specifications:
s	simple
dbl	two-handed
sm	simultaneous
alt	successive
arc	arc-shaped movement
mirr	mirroring
inv	inverse (back and forth movement)
crcl	circular

Abbreviations in section 3 and 4

PL	plural meaning
DISTR	distributive plural
ITER	iterative
REC	reciprocal

Non-manual markers (glossed above the main glosses)

br	brow raise
bht	backward head tilt

References

Aronoff, Mark, Irit Meir & Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81(2). 301–344.

Baker, Anne & Beppie van den Bogaerde. 2012. Communicative interaction. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, 489–512. Berlin: De Gruyter Mouton.

Barberà, Gemma & Patricia Cabredo Hofherr (eds.). 2018. *Sign Language & Linguistics: Special issue on Impersonal Reference in Sign Languages*.

Bauer, Anastasia. 2014. *The use of signing space in a shared sign language of Australia* (Sign Language Typology 5). Boston: De Gruyter Mouton.

Bergman, Brita & Östen Dahl. 1994. Ideophones in Sign Language? The place of reduplication in the tense-aspect system of Swedish Sign Language. In Carl Bache, Hans Basbøll & Carl-Erik Lindberg (eds.), *Tense, Aspect and Action*. Berlin: De Gruyter Mouton.

Brentari, Diane (ed.). 2010. *Sign Languages*. Cambridge: Cambridge University Press.

Burkova, Svetlana. 2015. The ways of expressing nominal plurality in the Russian sign language. *Sibirskij fonologicheskij zhournal* 2. 174–184.

Burkova, Svetlana & Elizaveta V. Filimonova. 2014. Reduplikatsija v russkom zhestovom yazyke (Reduplication in Russian Sign Language). *Russkij Yazyk v Nauchnom Osveshchenii* 28. 202–258.

Cormier, Kearsy, Adam Schembri & Bencie Woll. 2013. Pronouns and pointing in sign languages. *Lingua* 137. 230–247. https://doi.org/10.1016/j.lingua.2013.09.010.

Costello, Brendan. 2016. *Language and modality: Effects of the use of space in the agreement system of lengua de signos española (Spanish Sign Language)*. Amsterdam: University of Amsterdam Doctoral dissertation.

Daniel, Michael & Edith Moravcsik. 2013. The associative plural. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/36.

De Vos, Connie & Roland Pfau. 2015. Sign language typology: The contribution of rural sign languages. *Annual Review of Linguistics* 1(1). 265–288. https://doi.org/10.1146/annurev-linguist-030514-124958.

Filimonova, Elizaveta V. 2012. Sredstva vyraženija distributivnoj množestvennosti v russkom žestovom jazyke (Means of expressing distributive plurality in Russian Sign Language). In Olga V. Fedorova (ed.), *Russkij Zhestovij Jazyk. Pervaja Lingvisticheskaja Konferentsija. (Russian Sign Language. The first linguistic conference).*, 82–97. Moscow.

Filimonova, Elizaveta V. 2016. *Funktsionalno-semanticheskaja kategorija aspektualnosti v russkom zhestovom jazyke*. Moscow: MGU Doctoral dissertation.

Friedman, Lynn A. 1975. Space, time, and person reference in American Sign Language. *Language* 51(4). 940. https://doi.org/10.2307/412702.

Heyerick, Isabel, Mieke Van Braeckevelt, Jorn Rijckaert, Danny De Weerdt, Mieke Van Herreweghe & Myriam Vermeerbergen. 2011. *Meervoud in Vlaamse Gebarentaal*. Vlaams GebarentaalCentrum.

Johnston, Trevor A. & Adam Schembri. 2007. *Australian sign language (Auslan): an introduction to sign language linguistics*. Cambridge, UK ; New York: Cambridge University Press.

Kimmelman, Vadim. 2015. Quantifiers in RSL: distributivity and compositionality. In Peter Arkadiev, Ivan Kapitonov, Yury Lander, Ekaterina Rakhilina & Sergei Tatevosov (eds.), *DONVM SEMANTICVM*, 123–136. Moscow: LRC Publishers.

Kimmelman, Vadim. 2017. Quantifiers in Russian Sign Language. In Denis Paperno & Edward L. Keenan (eds.), *Handbook of Quantifiers in Natural Language: Volume II*, 803–855. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-44330-0_16.

Kimmelman, Vadim. 2018. Impersonal reference in Russian Sign Language. *Sign Language & Linguistics* 21(2). 204–231.

Klima, Edward & Ursula Bellugi. 1979. *The signs of language*. Cambridge, MA: Harvard University Press.

Kuhn, Jeremy & Valentina Aristodemo. 2017. Pluractionality, iconicity, and scope in French Sign Language. *Semantics and Pragmatics* 10(6). https://doi.org/10.3765/sp.10.6.

Liddell, Scott K. 2000. Indicating verbs and pronouns: pointing away from agreement. In Karen Emmorey & Harlan L. Lane (eds.), *The signs of language revisited: An anthology in honor of Ursula Bellugi and Edward Klima*, 303–320. Mahwah, NJ: Lawrence Erlbaum Associates.

Lillo-Martin, Diane. 2002. Where are all the modality effects? In Richard P. Meier, Kearsey Cormier & David Quinto-Pozos (eds.), *Modality in signed and spoken languages*, 241–262. Cambridge: Cambridge University Press.

Maroney, Elisa M. 2004. Aspect in ASL: A typological-functional analysis. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 30, 244–255.

McBurney, Susan L. 2002. Pronominal reference in signed and spoken language: are grammatical categories modality-dependent? In Richard P. Meier, Kearsey Cormier & David Quinto-Pozos (eds.), *Modality in signed and spoken languages*, 329–369. Cambridge: Cambridge University Press.

McBurney, S. L. 2004. *Referential Morphology in Signed Languages*. University of Washington Doctoral dissertation.

Meier, Richard P. 1990. Person deixis in ASL. In Susan Fischer & Patricia Siple (eds.), *Theoretical Issues in Sign Language Research, Volume 1: Linguistics*, 175–190. Chicago: University of Chicago Press.

Meir, Irit. 2012. Word classes and word formation. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language. An international handbook*, 77–112. Berlin: De Gruyter Mouton.

Meir, Irit & Wendy Sandler. 2008. *A language in space: the story of Israeli Sign Language*. New York, NY: Lawrence Erlbaum Associates.

Mel'čuk, Igor. 1993. *Cours de morphologie générale*. Montréal: Les Presses de l'Université de Montréal.

Miljan, Merilin. 2003. Number in Estonian Sign Language. *Trames* 7(3). 203–233.

Neidle, Carol & Joan Nash. 2012. The noun phrase. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, 265–292. Berlin: De Gruyter Mouton.

Nijhof, Sibylla & Inge Zwitserlood. 1999. Pluralization in Sign Language of the Netherlands (NGT). In *OTS Yearbook 1998–1999*, 58–78. UiL OTS.

Padden, Carol A. 1988. *Interaction of morphology and syntax in American Sign Language* (Outstanding Dissertations in Linguistics). New York, NY: Garland.

Perniss, Pamela, Robin L. Thompson & Gabriella Vigliocco. 2010. Iconicity as a General Property of Language: Evidence from Spoken and Signed Languages. *Frontiers in Psychology* 1. https://doi.org/10.3389/fpsyg.2010.00227.

Pfau, Roland & Josep Quer. 2010. Nonmanuals: their prosodic and grammatical roles. In Diane Brentari (ed.), *Sign Languages*, 381–402. Cambridge: Cambridge University Press.

Pfau, Roland & Markus Steinbach. 2021. Number in sign languages. In Patricia Cabredo Hofherr and Jenny Doetjes (eds.), The Oxford Handbook of Grammatical Number, 644–660. Oxford: Oxford University Press.

Pfau, Roland & Markus Steinbach. 2003. Optimal reciprocals in German Sign Language. *Sign Language & Linguistics* 6(1). 3–42. https://doi.org/10.1075/sll.6.1.03pfa.

Pfau, Roland & Markus Steinbach. 2006. Pluralization in sign and in speech: A cross-modal typological study. *Linguistic Typology* 10(2). https://doi.org/10.1515/LINGTY.2006.006.

Plungian, Vladimir. 2003. *Obshaja morfologija [General morphology]*. Moscow: Editorial URSS.

Reagan, Timothy. 2007. The times of our signs: aspect and aspectual markers in American sign language. *Southern African Linguistics and Applied Language Studies* 25(1). 17–26. https://doi.org/10.2989/16073610709486443.

Sandler, Wendy & Diane C. Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.

Schlenker, Philippe & Jonathan Lamberton. 2019. Iconic plurality. *Linguistics and Philosophy* 42(1). 45–108. https://doi.org/10.1007/s10988-018-9236-0.

Schmaling, Constanze. 2000. *Maganar Hannu, Language of the hands: A descriptive analysis of Hausa Sign Language*. Hamburg: Signum.

Stavans, Anat. 1996. One, two, or more: The expression of number in Israeli Sign Language. In William H. Edmondson & Ronnie B. Wilbur (eds.), *International Review of Sign Linguistics*, 95–114. Mahwah, NJ: Erlbaum.

Sutton-Spence, Rachel & Bencie Woll. 1999. *The linguistics of British Sign Language*. Cambridge: Cambridge University Press.

Tsay, Jane & James Myers. 2009. The morphology and phonology of Taiwan Sign Language. In James Tai & Jane Tsay (eds.), *Taiwan sign language and beyond*, 83–130. Chia-Yi: The Taiwan Institute for the Humanities.

Valli, Clayton & Ceil Lucas. 2000. *Linguistics of American Sign Language: An Introduction*. 3rd edn. Washington, D.C.: Gallaudet University Press.

Vermeerbergen, Myriam, Lorraine Leeson & Onno Crasborn (eds.). 2007. *Simultaneity in sign language: Form and Function*. Amsterdam: John Benjamins.

Wilbur, Ronnie B. 2010. The semantics–phonology interface. In Diane Brentari (ed.), *Sign Languages*, 355–380. Cambridge, MA: Cambridge University Press. https://doi.org/10.1017/CBO9780511712203.017.

Zeshan, Ulrike. 2000. *Sign Language in Indo-Pakistan: A description of a signed language*. Amsterdam: John Benjamins.

Zeshan, Ulrike, Cesar Ernesto Escobedo Delgado, Hasan Dikyuva, Sibaji Panda & Connie de Vos. 2013. Cardinal numerals in rural sign languages: Approaching cross-modal typology. *Linguistic Typology* 17(3). https://doi.org/10.1515/lity-2013-0019.

Zwitserlood, Inge. 2012. Classifiers. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, 158–186. Berlin: Signs of the time. Selected papers from TISLR 2004.

Zwitserlood, Inge, Pamela Perniss & Aslı Özyürek. 2012. An empirical investigation of expression of multiple entities in Turkish Sign Language (TİD): Considering the effects of modality. *Lingua* 122(14). 1636–1667. https://doi.org/10.1016/j.lingua.2012.08.010.

Paolo Acquaviva and Michael Daniel

23 Number in grammar: results and perspectives

1 Introduction

We outlined in chapter 1 the goals of this survey of number across a diverse sample of languages: to investigate the properties of number systems in some depth, while at the same time guaranteeing direct comparability between the analyses of systems that can be very different from each other. At the conclusion of this survey, it is appropriate to consider what picture emerges from it. We will structure our answer in three steps. First, in section 2, we briefly review the main results arising from the chapters in this volume, and we organize the typology of number values that emerge from them. Most of these considerations broadly confirm what is generally known (or assumed) about the expression, content, and variation space of number systems across natural languages. The next sections, from 3 to 7, discuss in greater depth and in a more analytical perspective some important themes, especially focusing on issues that can shed new light or contribute to current understanding of number systems. Finally, section 8 offers a wrap-up discussion and points to desiderata that arise from these studies.

For each language mentioned below, we provide genealogical affiliation as given by the contributor; for languages outside our sample, the affiliation is based on Glottolog 4.0 (Hammarström et al. 2021). Affiliation is provided when a language is mentioned for the first time, or when it has not been mentioned for a long time.

2 Number: diversity and uniformity

Certain features of a landscape can only be perceived at ground level, like the gradient of a street or the hedges screening the fields from view on a country lane. To see other features, one needs a higher vantage point; and for others yet, a map. In the same way, many significant aspects of the number systems described in the previous chapters emerge from comparing specific points across different chapters; others are best appreciated in the context of a whole chapter or of several whole chapters. Here we concentrate on aspects that emerge from all the foregoing chapters taken together.

2.1 Diversity of patterns

These studies evidence at the same time both a great diversity and a significant invariance. To fully appreciate the diversity of the number systems reviewed, it

should be remembered that the sample mostly included languages where number is prominent in grammar. It is within this class, very large but not encompassing the full typological spectrum, that the foregoing studies have illustrated substantially different ways to linguistically encode number information. One major distinction opposes systems where number has the characteristics of a robust morphological and morphosyntactic category (obligatory when applicable, without major gaps in the lexicon, marked on various lexical categories with a uniform range of values, underlying agreement) from those that express number-related information depending on lexical category (typically pronouns, less often on lexical nouns and verbs, and subject to the Animacy Hierarchy). The following languages illustrate the former subgroup: Arabic (Semitic), Kiowa (Kiowa-Tanoan), Lower Sepik languages, Nganasan (Uralic), Occitan (Romance branch of Indo-European), Slovenian (Slavic branch of Indo-European), Tswana (Southern Bantu) illustrate the former subgroup. Gooniyandi (Bunuban), Indonesian (Malayo-Polynesian branch of Austronesian), Japonic languages, and Mohawk (Iroquoian) illustrate the second, in very different ways. The very notion of part of speech, or word class, or lexical category, is not uniform, because while most languages clearly distinguish pronouns, nouns, and verbs as mutually exclusive lexical classes with fixed boundaries, some systems described in the book allow lexical stems to act as different categories according to the syntactic context (West Circassian, Mohawk, especially Indonesian; to place this type of variation in its typological context, see Bisang 2010). What is more, even the distinction between nominal and pronominal marking may call for a nuanced definition of what counts as "pronominal"; see the discussion of Standard Chinese (Sinitic branch of Sino-Tibetan) and Japonic in section 3.

The typological distinction between languages with "flexible" or "rigid" parts of speech is not the only major source of diversity in the way number is realized in grammatical systems. Unsurprisingly, our sample confirms that pronouns are the category that displays most number oppositions; but this does not mean that the expression of number is uniform across the lexical categories that it does apply to. It is common for pronominals to show more distinctions than nouns, and/or for number to be obligatory on the former but optional on the latter; but some systems additionally display deeper asymmetries. These are especially salient in those languages that display little or no number specifications on nouns but express the number information relative to arguments mainly by means of pronominal affixes (Marori (Trans-New Guinea), Mohawk), possibly in tandem with nominal phrasal clitics (Gooniyandi), in both cases clearly distinct from agreement affixes. The extent to which verbal number is developed as an autonomous category (as distinct from agreement) also varies greatly, with systems such as Karko (Nilo-Saharan) displaying a remarkable complexity, and others such as Russian Sign Language showing a significant overlap with the way number is expressed outside of verbs.

The range of possible number values also displays considerable variation, with singular, dual, and plural (possibly with oppositions between singular+dual vs. plu-

ral or singular vs. dual+plural, as in Marori), complemented by paucal, greater plural, along with overt markers signalling a distributive or a collective interpretation, in one case even overtly distinguishing genericity (Kakataibo, Panoan). In some cases, most dramatically in Kiowa, the exponence of number is complicated by the need to distinguish the value as a property of morphology, of morphosyntax, and of semantics; Kiowa in particular adds the twist of using an "inverse" marker which switches the value inherent in a base form. When distributivity is indicated, this is typically most prominent in verbs, where it may express number information indirectly by implication (scattering implying plurality), as in Mohawk and in Russian Sign Language.

In sum, our sample certainly evidences a great diversity in the way languages encode number information. Importantly, this is not just due to major typological differences between systems, but concerns the regularity of number marking across lexical categories and the range of number values, independently of typological similarity or dissimilarity.

2.2 Conceptual uniformity

In the face of all this, it may sound surprising to speak of a "significant invariance". But it takes little to realize that, behind this great variability, the manifold phenomenology of linguistic number is based on a few invariant parameters. All descriptions in this volume share the same overall structure based on the Questionnaire (chapter 1), with a detailed examination of the morphology of nominal, pronominal, and verbal number, followed by a discussion of the salient syntactic and semantic and pragmatic properties. Within this unified structural framework, the similarities in the way number systems are organized stand out all the more clearly. The opposition between singular and plural is indeed fundamental, even when it is modulated by the association of dual readings to one or the other value, or when 'singular' or 'plural' (or 'dual') are semantic and morphosyntactic but not morphological categories, like for instance in Kiowa (Kiowa-Tanoan), where exponents are not unambiguously singular, plural, or dual, but acquire their content by virtue of their paradigmatic distribution. The other values, which are not always really "minor" in the context of a grammatical system, go back to the established notions of paucal, greater plural, distributive, and collective. This means that these basic notions can be deployed as descriptive and analytical tools while doing justice to the full variability of the phenomena; number systems do not seem to be based on binary oppositions such as, say, having a cardinality of more or less than 10, or being odd-numbered or even-numbered, or on three-way oppositions like having 1 member, 5 members, or any other cardinality (cf. Number Hierarchy in Corbett 2000: 38–50). In addition, our survey also makes it clear that values like paucal or distributive or collective, not reducible to singular, dual, and plural, are anything but a secondary

aspect of the linguistic encapsulation of number. While they may be secondary in certain respects, they are part of the invariant conceptual categories necessary for understanding the organization of number across languages, especially when this is considered in its full extension, including verbal number. In sum, the foregoing studies confirm that the number values that have emerged in comparative studies are an appropriate conceptual vocabulary to account for a very diverse empirical landscape.

The validity of the basic number values established by comparative research goes hand in hand with the validity of some other results about these values. A dual necessarily presupposes the existence of a distinct plural, and in so far as a trial is distinguishable from numeral modification (see the discussion of Russian Sign Language in Kimmelman et al., this volume), this presupposes a dual. Likewise, outside of the closed system of pronouns, the singular very often lacks a distinctive marking, not only where number marking is optional on nouns; and more generally, if either the singular or the plural lacks morphological exponence, it will never be the plural. This is true also where an "inverse" marker switches the value that is inherent in a given choice of noun (as in Kiowa), or where an affix turns inherently plural nouns into singulars (as in Arabic, Semitic, and Karko, Nilo-Saharan). In both cases, this applies to a specific subset of the nominal lexicon without precluding the existence of overtly marked plurals in the language so the generalization holds that no language uniformly needs overt marking for the singular but not for the plural.

As for the plural itself, the canonical "additive" interpretation of a pluralized noun N as 'more than one (or two) entities each of which is an N' is of course central, but a plural denotation based on N may correspond to an associative interpretation, especially with proper names ('N and pragmatically or contextually related individuals'), or "similative" ('N and similar things'); see for instance Shimoji's discussion of Japonic languages (this volume, 4.1.2, 4.3). Also related to the interpretive range of plurality is the distinction between inclusive and exclusive readings on pronouns, which will likewise be taken up more in detail below.

2.3 The number category and part of speech typology

Mention of pronouns brings us to the issue of the variability of number marking across lexical categories. We have already mentioned that our sample confirms that pronouns tend to display a greater differentiation than nouns in terms of number marking, and that lexical categories often vary in the range of values that they can be marked for. This holds good generally, even though it should be noted that a language like Pilagá (Guaicuruan), reported by Krasnoukhova (this volume), shows that there can be systems with more distinctions on nouns than on pronouns.

However, what matters is not so much the greater or smaller number of distinctions but rather the fact that such differences may signal a deeper difference in the

way number is organized, depending on the lexical category. The possible misalignment between pronouns and nouns is a theme that re-emerges in various chapters, especially those related to clusivity (see section 3); the asymmetry between (pro)-nominal and verbal number is deeper, and goes back to a more fundamental difference between the linguistic characterization of arguments and of events (see section 6).

Another well-known parameter that underlies the variability of number systems is the Animacy Hierarchy, which is directly related to the variability across lexical categories but is also sensitive to other variables. Overall, the data discussed in the foregoing chapters confirm that grammatical number oppositions occupy a continuous segment along the axis defined by the Animacy Hierarchy (Smith-Stark 1974, Corbett 2000: 56), so that for instance a differentiation on inanimate nouns presupposes one on animate nouns, this presupposes one on third-person pronouns, and this in turn presupposes one on first- and second-person pronouns (for one comparative overview, see Krasnoukhova, this volume, 2.2.1, 2.3.1). However, the facts are not as uniform as this generalization would make us expect. For instance, Yimas (Lower Sepik) independent pronouns show a four-way distinction in the number paradigm of third-person bound forms, but, with one exception, only a three-way distinction in the first and second (Foley, this volume, 2.1, tables 8 and 9), which are higher according to the hierarchy and are expected to display more (rather, not less) contrasts; and in Marori (Trans-New Guinea), verbal number displays a three-way distinction in the second person, as opposed to a two-way distinction in the first (Arka and Dalrymple, this volume, 2.4). Facts such as these may problematize the conception of pronominals as the upper section of the Animacy Hierarchy (understood in terms of sensitivity to number distinctions), as discussed below in section 5.

3 Pronominal number

Pronominal reference is central to human cognition because of its importance for social interaction. Notwithstanding the great variation in how the distinctions in the domain of person may combine to shape a language-specific design, some of these distinctions are good candidates for language universals. Much work in the literature has been devoted to studying the observed configurations with a view to revealing which of them are frequent, which are rare, and which are apparently impossible, with the aim of arriving at a metalanguage to describe person interwoven with number (Ingram 1978, Sokolovskaja 1980, Greenberg 1988, Cysouw 2001, Harbour 2016).

Pronominal reference is richly covered in descriptive grammars and has a long record of typological research, from Forchheimer (1955) to Ingram (1978), Cysouw (2001), Bhat (2003), and Siewierska (2004), to cite just a few examples. Number distinctions in pronouns are prominent in these discussions and this interest seems

to hold across frameworks. It is thus no wonder that this volume, whose main purpose is to collect evidence for empirical diversity, does not bring groundbreaking insights with respect to pronominal number. Some of the languages covered show relatively rare patterns, such as pronouns sharing plural morphology with nouns in Japonic languages (Shimoji, this volume), Kakataibo, Panoan (Zariquiey, this volume) and more generally in South America (Krasnoukhova, this volume, 2.2.1); and optional pronominal number discussed by Krasnoukhova (this volume) and Gil (this volume, 2.2). Our sample also includes the pronominal system of Gooniyandi (Bunuban), which is quasi-unique according to McGregor (1996); see also McGregor (this volume). These unusual cases highlight the diversity covered in our volume but we are not presenting them as providing novel evidence as none of the systems described in the volume is previously unattested.

The idea that pronominal number is a category different from nominal number is now received knowledge robustly planted in linguistic textbooks. This is clear in conceptual terms, the pronominal number being almost necessarily referentially heterogeneous, labeled 'approximative' (in the sense of not all elements of the set being the same) in Jespersen (1924) or 'representative' (in the sense of one element of the set 'standing for' all other elements) in Barulin (1980), or, in more recent terminology, designating heterogeneous sets including a prominent *focal* referent and secondary referents (Moravcsik 2003; cf. also *focus* and *spectrum* in Austerlitz 1959). This is opposed to nominal number which is (typically) homogeneous, or 'additive' (in the sense of several denotata of the same nominal being 'added up' to form a set) in Barulin (1980). The difference is also clear in terms of exponence, with a very visible cross-linguistic preference of number distinctions expressed on pronouns more often than on nouns (for example, Gil, this volume, Krasnoukhova, this volume) together with a dispreference against pronouns and nouns sharing formal expression of plurality (see Daniel 2013 for cross-linguistic counts, but also Krasnoukhova, this volume, who shows that this pattern is not so rare in South America).

These known properties of pronominal plurals are so widely discussed in the literature that they may mask subtler differences between pronominal and nominal number. As one example, the dual of personal pronouns does not show the same attraction towards cohesion, i.e. designation of sets of elements that naturally occur together, as the nominal dual (Plank 1989, and section 4.6 below). To be precise, the effects of cohesion are sometimes also observed in pronominal plurals as discussed by Corbett (2000: 26–30) on Sursurunga or mentioned more generally for Western Oceanic languages by Dalle Ceste (ms.). Sometimes these cohesive plurals are paradigmatically restricted, as when the distinction between regular plural pronouns and special forms to designate a circumscribed group such as 'community' or 'family' is restricted to the possessive forms of personal pronouns in Khinalug, East Caucasian (Kibrik et al. 1972).

Viewed differently, one could also say that a certain pragmatic cohesion is built into pronominal plurality anyway because it is a contextualized relation that allows

the interlocutor to reconstruct the rest of the group referred to by plural personal pronouns 'we' and 'you.pl' by their association with the speaker or the addressee, respectively. This led various scholars to describe the category of number in personal pronouns as *group* (Moravcsik 1994, Cysouw 2001, Siewierska 2004: 86) or *associative* (Corbett 2000: 103, Cysouw 2001: 66; Vassilieva 2005; Kiparsky 2013), conceptually aligning plurals of personal pronouns with nominal associative plurals ('X and his/her family/friend/associates'). However, with the rare exception of group pronouns as attested in e.g. Oceanic, the exact nature of this relation is unspecified. As opposed to nominal associative plurals ('X and associates'), the meaning of a personal plural pronoun is probably better described simply as 'a group that includes the speaker' (or 'the addressee'), the contextual interpretation of the grouphood left fully to the addressee's ability to reconstruct the composition of the group based on the speech situation. It is this lack of specification of grouphood in pronominal plurals that has been suggested as a functional motivation for *inclusory constructions*, also attested in various languages in this volume, as a means of making the reference of a plural pronoun more directly accessible to the addressee (Lichtenberk 2000). Therefore, we prefer to use *heterogeneous* plurality as a cover term for pronominal and associative plurals in the sense that, in plural personal pronouns, the speaker is included into a group of non-speakers, or the addressee is included into a group of non-addressees (cf. Mauri and Sansò 2022).

On the whole, distinctions in cohesion seem less relevant to the typology of pronominal number than of nominal duals because pronominal cohesion is either universal and thus in a sense trivial, with any 'we' and 'you.pl' suggesting a certain pragmatically established group, or else marginal, dedicated *group* pronouns being crosslinguistically rare. Together with the common observation that trials and the more disputed quadrals (Corbett 2000: 26–30, Dalle Ceste, ms., and the data from Russian Sign Language in Kimmelman et al., this volume) are only observed in pronouns and not in nouns, this seems to be an indication that pronominal number is more prone to encode pure cardinalities (cardinalities '2', '3' and possibly '4') than nominal number (only cardinality '2', sometimes further restricted to cohesive sets).

But even in such a well-covered domain, a typologically informed description may contribute to a better understanding of an individual system. The Nivkh (isolate) pronoun *meŋ* (and its correspondances in other dialects) has been traditionally described simply as a first person dual. A system with a stand-alone dual pronoun in the first person and an inclusive pronoun in the plural is strongly reminiscent of Greenberg's substitute for pronominal number, the "minimal-augmented" category, where the dual inclusive is aligned with singular pronouns (Greenberg 1988, and his predecessors such as Thomas 1955, Conklin 1962, Boxwell 1967, McKay 1978; also McGregor 1989, Greenberg 1989, McKay 1990; and overviews in Corbett 2000: 166–169, Cysouw 2001, 3.6.5 and 4.5.2, Siewierska 2004: 82–88). This analysis requires that the dual first person pronoun be dedicated to inclusive reference, something not previously observed in the grammars of Nivkh (e.g. Panfilov 1962). Now, Gruzde-

va (1998, also this volume) confirms that *meŋ* seems to be only attested with inclusive reference, so that the whole system may be re-interpreted. Fig. 1a shows the alternative interpretations of pronominal number in the systems of Nivkh, illustrated with the forms of the West Sakhalin dialect. Systems involving a 'trial inclusive', not represented in the languages in the volume, are shown in Fig. 1b as a comparison of two analyses of pronominal systems in Rembarrnga (Gunwinyguan) and Djeebbana (Maningrida) as reported in McKay (1978), illustrated with the forms of possessive pronouns of Djeebbana.

		Sg	Du	Pl		Minimal	Augmented
1st	Excl	*ɲi* 'I'	NA	*ɲəŋ* 'I and other(s)'	1st	*ɲi* 'I'	*ɲəŋ* 'I and other(s)'
	Incl	NA	*meki* 'you.sg and I'	*mer* 'you.sg and I and other(s)'	Incl	*meki* 'you.sg and I'	*mer* 'you.sg and I and other(s)'
2nd		*cʰi* 'you.sg'	NA	*cʰəŋ* 'you.sg and other(s)'	2nd	*cʰi* 'you.sg'	*cʰəŋ* 'you.sg and other(s)'

Fig. 1a: West Sakhalin Nivkh pronouns in the conventional representation (left) and in a Greenbergian perspective (right).

		Sg	Du	Tri	Pl		Minimal	Unit augmented	Augmented
1st	Excl	*njabba* 'I'	*njaddana* 'I and another person'	NA	*njaddabirra* 'I and others'	1st	*njabba* 'I'	*njadda-na* 'I and another person'	*njadda-birra* 'I and others'
	Incl	NA	*ngadda* 'you.sg and I'	*ngaddana* 'you.sg and I and another person'	*ngaddabirra* 'you.sg and I and others'	Incl	*ngadda* 'you.sg and I'	*ngadda-na* 'you.sg and I and another person'	*ngadda-birra* 'you.sg and I and others'
2nd		*ngka* 'you.sg'	*naddana* 'you.sg and another person'	NA	*naddabirra* 'you.sg and others'	2nd	*ngka* 'you.sg'	*nadda-na* 'you.sg and another person'	*nadda-birra* 'you.sg and others'

Fig. 1b: Djeebbana (Maningrida) possessive pronouns in the conventional representation (left) and in a Greenbergian perspective (right).

The privileged status of the dual inclusive is highlighted by the universal suggested in Moravcsik (1978: 352) and discussed in Plank (1989), which may be read as follows: no dual exclusive exists without a dual inclusive, while the opposite is possible, as shown in Fig. 1a and 1b. A diagnostic feature of this kind of pronominal systems is that the plural (augmented) inclusive pronoun is different from the other plural pronouns in terms of cardinality. For example, the plural inclusive pronoun *mer* in Nivkh has the minimum cardinality of three while the other Nivkh plural pronouns have the minimum cardinality of two. Accordingly, the plural inclusive pronoun *ngaddabirra* in Djeebbana has the minimum cardinality of four while the

other Djeebbana plural pronouns have the minimum cardinality of three. This cannot be explained in terms of the traditional analysis (the left panes in Fig. 1) but is accounted for by the Greenbergian one (the right panes in Fig. 1). Indeed, unlike the other persons, the inclusive minimal category consists of two referents, so that the Nivkh augmented inclusive *mer* minimally refers to a group of three people, and the Djeebbana augmented inclusive *ngadda-birra* minimally refers to a group of four people (because the inclusive reference to a group of three is expressed by a dedicated unit augmented pronoun *ngadda-na*).

In some languages – including in Rembarrnga and Djeebbana – the Greenbergian analysis is further supported by the morphological expression of these categories because it allows a meaningful analysis of the morphological structure of the pronominal forms (Greenberg's *transparent* subtype of minimal-augmented systems). Thus, for Djeebbana in Fig. 1b the suffixes *-na* and *-birra* can be related to the expression of the values *unit augmented* and *augmented*. But even in the *opaque* subtype represented by Nivkh, Greenberg's analysis is preferable in that it eliminates the otherwise arbitrary gaps that appear in the left-hand pane, as in Fig. 1a. We will be saying that the dual inclusive is *paradigmatically aligned* with singular pronouns of the other persons in languages like Nivkh and Djeebbana; and that the trial inclusive is aligned with dual pronouns of the other persons in languages like Djeebbana. To throw the special nature of the minimal augmented systems in relief, Fig. 2 compares them to another, more familiar type featuring other dual pronouns with which the dual inclusive is aligned, as in Kove, Oceanic branch of Austronesian (Sato 2013 via Dalle Ceste, ms.; the special series of group pronouns for cohesive sets is omitted).

		Sg	Du	Pl		Minimal	Unit augmented	Augmented
1st	**Excl**	*yau* 'I'	*yahua* 'I and another person'	*yai* 'I and others'	**1st**	*njabba* 'I'	*njadda-na* 'I and another person'	*njadda-birra* 'I and others'
	Incl	NA	*tahua* 'you.sg and I'	*taita* 'you.sg and I and other(s)'	**Incl**	*ngadda* 'you.sg and I'	*ngadda-na* 'you.sg and I and another person'	*ngadda-birra* 'you.sg and I and others'
2nd		*veao* 'you.sg'	*amihua* 'you.sg and another person'	*amiu* 'you.sg and others'	**2nd**	*ngka* 'you.sg'	*nadda-na* 'you.sg and another person'	*nadda-birra* 'you.sg and others'
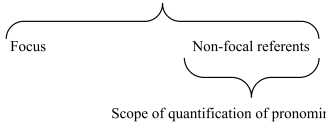
Fig. 2: Alignments of the inclusive pronouns in Kove, Oceanic (left, traditional number categories) and in Djeebbana, Maningrida (right, Greenbergian perspective).

Unlike Fig. 1 where two different approaches to the same system are compared, Fig. 2 compares two different types of systems. Fig. 2, left pane, suggests that, for Kove, it is the traditional analysis in terms of cardinality values that works better – note the morphology shared by the pronouns with cardinality two. The gap in the table is inevitable because an inclusive pronoun with cardinality one is logically

impossible. For Djeebbana in the right pane, on the other hand, it is the Greenbergian analysis that works better in terms of both morphology and absence of paradigmatic gaps. In other words, what this figure shows is the existence of empirical differences between types of systems rather than merely theoretical differences in analyses.

Greenberg's minimal value corresponds to the focus in our terminology, assuming that the focus of the inclusive pronoun in Nivkh and Djeebbana is a dyad that includes both the speaker and the addressee. The difference between the two types of systems in Fig. 2 is not that one is and the other is not related to quantification. Both are means of quantification of heterogeneous sets, which is exactly what pronominal number is. The difference is what is quantified. On the left in Fig. 2, pronominal number counts the cardinality of all referents of the pronoun, including its focus (*singular* for one, *dual* for two, *plural* for more). These systems are *cardinality-based*. On the right, Greenberg's *minimal*, *unit augmented* and *augmented* categories do not count the cardinality of the sets denoted by pronouns but only that of non-focal referents added to the focus (*minimal* for none, *unit augmented* for one, *augmented* for more). This accounts for the special effect of the increase in the minimal cardinality of the inclusive pronouns as compared to the pronouns of other persons discussed above. The difference in the scope of quantification is shown in Fig. 3.

Cardinality-based system of pronominal number: e.g. Karko, Ryukyuan, Riau Indonesian, Kupang Malay, and Kove	Scope of quantification of pronominal number
	Focus Non-focal referents
Minimal-augmented system of pronominal number: e.g. Nivkh and Rembarrnga	Scope of quantification of pronominal number

Fig. 3: Comparison of two types of pronominal number: Scope over the referents of the pronoun.

With the adjustment for the scope of quantification, the two types become more similar than the use of the conventional term *pronominal number* as opposed to Greenberg's dedicated category *minimal-augmented* may suggest. Pronominal number is inevitably heterogeneous (see also section 5), so it must not come as a surprise that some systems show additional sensitivity in that they avoid counting the more prominent focal referents together with the rest of the referents of the pronoun, the 'others'. As Corbett (2000: 169) notices, the oppositions of singular vs. plural and minimal vs. augmented are but two different instantiations of pronominal number, the category that linguistically encodes the difference between the forms referring to the focus, *singular* or *minimal*, and the forms referring to groups including the focus, *plural* or *augmented*, respectively, depending on the type of the system. However, as we are now going to show, spelling out the difference between cardinality-based and minimal-augmented systems, i.e. between systems counting vs. not-

counting the focal referents has important implications for modeling the category of inclusive and of pronominal number in general.

It is noteworthy that the distinction between minimal-augmented and cardinality-based systems arguably matches the two of Humboldt's three types of the dual, described by Plank (1989: 313) as "languages whose notion of the dual derives from the opposition between the speaker and the person spoken to" and "languages whose dual is not inspired by any particular kind of dyad but is based on the abstract and general notion of duality as such", respectively. Humboldt's third type are cohesive duals in "languages which obtain this grammatical form from the perception of objects naturally occurring in pairs", to be discussed in section 4.5 below.

So far we have faithfully followed Greenberg's seminal argumentation that suggested that, in minimal-augmented systems, the inclusive is a separate person. Our next step is extending the view of the inclusive as a separate person to all systems, including those counting the totality of referents. To do so, let's step back and try to embrace the traditional view of the inclusive as a first person, persistent in typology notwithstanding the ongoing discussion of inclusives and the evidence from minimal-augmented systems.

The traditional logic is as follows. The person of the plural pronoun is defined as the speech act role of its focus. Thus, English *we* is a first person pronoun because its focus is the speaker. Considering the inclusive a first person pronoun means assuming that the speaker is its one and only focal referent, the difference between the inclusive and the exclusive forms being delegated to an additional distinction in terms of involvement of one special non-focal referent, namely, the addressee (i.e., to the category of *clusivity*). Under the conventional definition of non-singular personal pronouns as heterogeneous, i.e. a group including the focus and additional referents, this entails that the inclusive pronoun, including the dual inclusive 'you.sg and I', is necessarily a non-singular pronoun. Indeed, it includes (at least) one referent in addition to the speaker, its focus. Assuming that the inclusive is a first person pronoun thus logically leads to the conclusion that the inclusive is a plural pronoun: *inclusive is a first person pronoun* $\implies$ *inclusive is a plural pronoun*.

In minimal-augmented systems, the logic has to be changed. This is because the inclusive pronoun 'speaker and addressee' is paradigmatically aligned with singular pronouns, whether purely in terms of the configuration of the slots (opaque systems) or also in terms of the morphological expression of number (transparent systems); cf. Fig. 1a and b). In these cases the inclusive is not a first person pronoun because in these languages 'speaker and addressee' cannot count as a plural pronoun and consequently inclusive cannot count as a subtype of the first person plural: *inclusive expressing 'speaker and addressee' is not a plural pronoun* $\implies$ *inclusive is not a first person pronoun*.

If, as argued above, the focus of the inclusive is not the speaker alone but both the speaker and the addressee (Greenberg's *minimal* inclusive) in minimal-augmented systems, what is the evidence that it is not also the case in the languages where

the inclusive is not aligned with singular pronouns? In all languages that have it, a dedicated inclusive pronoun necessarily includes both the speaker and the addressee. The conventional classification of the inclusive as a first person implicitly prioritizes the speaker over the addressee. However, there is no empirical evidence that the speaker has referential salience with respect to the addressee in the languages that have a pronoun dedicated to inclusive reference, whichever type of pronominal number they have according to Fig. 3. In particular, Daniel (2005) suggests that, crosslinguistically, there are no indications of a stronger formal association of the inclusive with the first or with the second person; in most languages this is an independent pronominal root (but see Pertsova, ms., for a typological overview of a large sample and evidence against this claim).

Classifying the inclusive as a first person reflects the perspective of the languages lacking the inclusive pronoun: it is in languages like English that any reference to groups including the speaker, whether inclusive or exclusive of the addressee, is performed by the same first person plural pronoun, with the focus invariably being the speaker. In these pronominal systems, the speaker is indeed prioritized over the addressee (Speaker > Addressee). For the discussion of the relation between the category of inclusive and person hierarchies see Zwicky (1977) and most articulately Plank (1985), who argues for the person hierarchy Speaker = Addressee as a functional explanation of the presence of the inclusive pronoun. This version of person hierarchy amounts to considering the dyad 'speaker and addressee' as the focus of all inclusive pronouns; see also Dixon (1994: 90) and Daniel (2005) for a review of the arguments, a model and some evidence from honorific uses of the inclusive pronouns.

Yet, outside languages like Nivkh, the inclusive continues to be viewed as a first person category. This is not done on empirical grounds but because these systems do not provide evidence to the contrary. It seems that interpreting the inclusive as a first person plural pronoun is effectively taken as the null hypothesis that does not need arguments to support it but needs arguments to reject it. Because the only generally accepted argument is the alignment of the dual inclusive with singular pronouns, a two-way logical connection is established between the two stipulations. Not only is the inclusive considered a non-first person pronoun if the expression of 'speaker and addressee' is aligned with singular pronouns of other persons, but also conversely, if the expression of 'speaker and addressee' is aligned with dual pronouns of other persons as in Kove in Fig. 2, or if it is not attested in the language at all as a separate pronoun and there is only a general inclusive pronoun covering both the meaning 'speaker and addressee' and the meaning 'speaker and addressee and other(s)' as in Riau Indonesian (Malayo-Polynesian; see Gil, this volume), the inclusive is considered a first person non-singular pronoun: *inclusive is a first person pronoun ⟺ the inclusive expressing 'speaker and addressee' is a plural (non-singular) pronoun.*

We argue that this last mutual implication does not hold. As explained above, the singular or non-singular alignment of the dual inclusive may be linked not to

its person affiliation but to which of the two types of pronominal number the language has. In this logic, the inclusive does not need to be a first person pronoun in order to be a plural pronoun. Fig. 4 classifies all inclusive pronouns as a separate person and explains differences between them in terms of the type of pronominal number.

	Focus	Counting only non-focal referents		Counting all referents				
		Nivkh		Kove			Riau Indonesian	
		Min	Aug	Sg	Du	Pl	Sg	Pl
1st	**speaker**	'I'	'I and other(s)'	'I'	'I and another person'	'I and others'	'I'	'I and other(s)'
Incl	**speaker and addressee**	'you.sg and I'	'you.sg, I and other(s)'	NA	'you.sg and I'	'you.sg, I and other(s)'	NA	'you.sg and I' or 'you.sg, I and other(s)'
2nd	**addressee**	'you.sg'	'you.sg and other(s)'	'you.sg'	'you.sg and another person'	'you.sg and others'	'you.sg'	'you.sg and other(s)'

Fig. 4: Inclusive across pronominal systems.

In Fig. 4, the difference between Nivkh, on the one hand, and Kove and Riau Indonesian, on the other, is not in terms of person categories (cf. Fig. 2 above) but in terms of what the respective categories of pronominal number count. For Kove and Riau Indonesian, gaps in the paradigm appear again; but, unlike the gaps in Fig. 1, they are fully predictable. Under the rules applied to counting in these systems, 'you.sg and I', which contains two referents, necessarily counts as a non-singular category, so, for the inclusive person, the singular is unavailable. Further, there is a difference between Kove and Riau Indonesian. In Kove, the pronominal number has the dual value, so 'you.sg and I' is a pronoun separate from 'you.sg, I and other(s)'. But Riau Indonesian has no dual category in pronouns. As a result, two different inclusive meanings, 'you.sg and I' and 'you.sg, I and other(s)' are merged in one form. This happens because these systems lack the dual and are only capable of distinguishing between singular and non-singular reference. In all persons, the number differences in pronouns distinguish between reference to the focal referent vs. reference to a group including the focal referent. But the focus of the inclusive is itself a group. As both the inclusive that denotes only the speaker and the addressee and the inclusive that denotes both the speaker and the addressee and non-focal referent(s) refer to more than one person, in Riau Indonesian they share the same exponence.

This resolves another issue that so far was unattended to. On the one hand, we argue that the inclusive is a separate person with the focus 'speaker and addressee'. On the other hand, we argue that pronominal plurals are heterogeneous, consisting of the focus and non-focal referents. In the case of the inclusive person, number

opposition is therefore expected to distinguish between reference to the speaker and the addressee (focus only) and reference to a group that, in addition to the speaker and addressee, also includes one or more non-focal referents. Yet, in many languages featuring the inclusive pronoun, this difference is not expressed, so that the inclusive pronoun seems to be unspecified for number even if the category of number is obligatory in other pronouns. The solution is that, in these languages, the content of the category of number is the cardinality of the whole group. In terms of such cardinality, both 'speaker and addressee' and 'speaker, addressee and other(s)' include more than one referent and represent the same number value, the inclusive being a kind of plurale tantum. On the other hand, minimal-augmented systems such as those on the left in Fig. 4 count the cardinality of non-focal referents, so that 'speaker and addressee' and 'speaker and addressee and other(s)' are necessarily formally distinct. In both cases, the heterogeneity of the plural is not itself the content of the category of number but a logical consequence of the fact that there cannot be more than one speaker or more than one addressee (see below in section 5 on the uniqueness presumption).

Instead of arguing – in what appears to be Greenberg's wake – that it is counting only non-focal referents that makes the dual inclusive a singular value of a separate person in minimal-augmented systems, we reverse the logic. We argue that it is counting all referents – both focal and non-focal ones – that makes the inclusive pronoun necessarily a non-singular category even if it only refers to the focus. In the systems that lack the dual value, this leads to a merger of the meanings 'speaker and addressee' and 'speaker, addressee and other(s)' in one pronoun. But just as in minimal-augmented systems, this plural pronoun has the speaker and the addressee as its focus and may not be viewed as a subcategory of the first person non-singular.

To sum up, considering different ways of conceptualizing pronominal number leads us to the suggestion that all pronominal systems featuring a pronoun or pronouns dedicated to inclusive reference are different from the systems lacking such pronoun(s) in that they have an extra person value, the *inclusive person* (focus equals 'speaker and addressee') rather than in that they split first person reference into two subtypes, inclusive and exclusive. What we tried above was not so much to suggest a new analysis (our approach to inclusives goes back to Plank 1985) but to show why the still surviving conventional model of the subcategorization of the first person into exclusive and inclusive is implausible on functional grounds, and connect this to the concept of pronominal plurality. Note that a similar conclusion has been reached in formal models of pronominal reference by decomposing person values into features [±Speaker] and [±Addressee] (see McKaughan 1959, Conklin 1962, followed by many formally-oriented studies, e.g. Nevins 2007). Not so willingly we suggest disposing of the recent, much celebrated and indeed elegant term *clusivity* (Filimonova 2005). There seem to be no such linguistic category as clusivity, and thus no inclusive or exclusive 'we'.

Obviously, many questions remain open. For example, at first glance, the difference between the two types of pronominal number is limited to the languages with the inclusive because so far it was by the position of the inclusive in the pronominal system that we decided what was counted. However, independently of the presence or absence of the inclusive, in languages where the plural personal pronouns are produced by applying nominal plural marking to the singular pronoun it is plausible to stipulate that pronominal number counts all referents just as nominal number does. In principle, this makes the discussion above falsifiable: we do not expect to find languages with minimal-augmented systems where personal pronouns and nouns have shared exponence of the category of number (unless, in nouns, this is a dedicated associative plural marker which is not used for the regular, additive plurals, and thus is compatible with the augmented 'focus and other(s)' interpretation). As a second example of an open question, consider the pronominal system of another language in the volume, Eastern Dan. This language has a dedicated dual inclusive pronoun as well as a plural inclusive, apparently replicating the minimal-augmented pattern of Nivkh as discussed above. However, Vydrin (this volume, 2.2.2, and note 3) observes that Eastern Dan deviates from the minimal-inclusive type in that its plural inclusive forms may also have dual inclusive reference, and in this function seem to be in free variation with the dedicated dual inclusive pronoun. The minimal cardinality of all plural pronouns is thus two, so that, despite appearances, pronominal number in Eastern Dan is of the cardinality-based type. The presence of a dedicated dual inclusive should be explained differently, probably by diachrony. As one speculative scenario, the original minimal-augmented system may be on its way to the future cardinality-based system by supplementing the uses of the former minimal – now dual – inclusive by the former augmented – now plural – inclusive; note that, according to Vydrin (personal communication), such uses only appear in natural texts, not in elicitations, which possibly indicates a change in progress. How common is this situation cross-linguistically? Distinguishing such cases from true minimal-augmented systems is also a descriptive challenge, as is always the case with free variation. See Cysouw (2001) for more on diverse implementations of minimal-augmented systems.

4 Nominal number

One of the dangers inherent in comparing the number values available in different languages lies in the implicit assumption that the values listed for any one language are all on a par. This assumption may be wrong, and not simply because values might differ in terms of frequency of use, or because the range of available values is not typically the same between nominals and verbs, and may also change between nouns and pronouns, and possibly adjectives. The point is rather that the number values available in a language can differ from each other in the type of role

they have in the grammatical system. The specific phenomena that motivate calling a range of values "minor" (Corbett 2000: 95–101), or viewing it as an extra qualification of size or numerical quantity, additional to what the core number values express, vary from language to language. Here we will briefly review some of the relevant phenomena that emerge from the chapters in this book.

4.1 Countability

The lexical distribution and applicability of number is obviously linked to the broad theme of countability, and, given that countability is related to individuation, also to the Animacy Hierarchy (see Corbett 2000, Chapter 3). In this respect, too, the studies in this book confirm the results of the most recent scholarship – which do not always match the traditional view. According to the received view, mass nouns are overwhelmingly morphologically singular, like English *blood*, *footwear*, or *fun*, except for plural-only cases like *suds*, *fumes*, or *dregs*; the pluralization of a singular mass noun while keeping its mass interpretation is supposed to be excluded, so that *a few wines* requires a re-categorization as a count noun, typically referring to sorts or concrete standard-sized portions. We called this the received view, because in fact already Allan (1980) made it clear that the morphosyntactic tests for countability do not always align perfectly (so, *three clothes* is very odd but *a few clothes* is not), and plurality is certainly compatible with a mass reading. Thanks to the great development of studies on countability in a comparative perspective (see Massam 2012 for one major illustration), we now know that the seemingly clear match of semantic and morphosemantic properties which gave rise to the very notion of 'mass noun' is much less systematic than it appeared (even in English), and that some systems make it very problematic, or impossible, to define mass nouns as a grammatical phenomenon. But even when this is possible, mentions of pluralized mass nouns are frequent in the foregoing chapters, and although they most usually denote types or distributed instances, a truly mass reading for such pluralizations is attested, especially in an intensive interpretation of abundance, as in Arabic (Semitic), see Fassi Fehri, this volume; this was apparently possible in older varieties of Nivkh (isolate) too, see Gruzdeva, this volume, 2.3.2.1; see also the examples from Lezgian (East Caucasian), Fula (Atlantic-Congo), and Koryak (Chukotko-Kamchatkan) in Corbett (2000: 238). We may add that in Dargwa languages (East Caucasian), many mass nouns are morphologically singular and may be morphologically pluralized, but show plural agreement irrespective of their morphological number; while in Archi (also East Caucasian), mass nouns show elements of plural inflection. Finally, it is important to flag the new perspective opened by the recent study of Tsiakmakis et al. (2021), which suggests that plurality on Greek (Indo-European) mass nouns has an expressive function and conveys a negative speaker attitude; how much this pragmatic perspective can help account for the use of plurality on mass in other languages is an open, and very interesting, question.

The overly schematic equation of "mass syntax" with "no number differentiation" has traditionally been reinforced by the traditional distinction between languages with number and languages with classifiers. This schematic distinction contrasts languages that make use of grammatical number and languages that make use of classifiers in order to represent grammatically the interpretation of a nominal as a countable bounded individual, as a mass, and as an unbounded plurality of individuals. Number and classifiers would thus be complementary grammatical strategies to linguistically express the "division of reference", to use Borer's (2005) terminology which in turn echoes Quine (1960). Approaches diverge on whether classifiers and number are alternative ways to express the same countability opposition through grammar, or whether countability can only be a grammatical property of nouns in number-based systems, while classifier-based systems treat all nouns as mass-like (see Krifka 1995, Chierchia 1998, Borer 2005, Massam 2012, Zhang 2013); but both hinge on a view of classifiers and number as mutually complementary. That classifiers exclude number is indeed a well-documented typological implication, but importantly it refers to *obligatory* classifiers and to *obligatory* number marking on nouns; this is the so-called Greenberg-Sanches-Slobin generalization, carefully discussed by Doetjes (2012) and more recently in Tang and Her (2019). Even so, Aikhenvald (2003: 249) notes several exceptions to this restricted version of the complementarity generalization. What is more, an increasing body of literature about numeral classifiers links the presence of classifiers to properties of the numerals themselves, rather than to any typologically relevant distinction in the properties of nouns, especially in the light of the observation that some languages require classifiers with low numerals but not other numerals (see Gil 2013, Bale and Coon 2014). It is against the background of these more nuanced results that the data emerging from our sample should be assessed.

Not many languages make use of classifiers, and in line with expectations, the issue of how classifiers interact with number only arises in a few cases. Japonic languages and Eastern Dan (Mande) have both classifiers and number in their grammatical systems (see Shimoji, this volume, and Vydrin, this volume), but they do not seem to co-occur. However, classifiers can co-occur with plurally-marked nouns in the same structure in some South American languages (see Krasnoukhova, this volume, 2.3.6); the same is reported as a limited innovation for Nivkh (Gruzdeva, this volume, 3.2, see discussion of example (102)), where classifiers are only obligatory up to the number 5. While the co-occurrence of plural marking and numerical classifier is not by itself a counterexample to the strictest version of the Greenberg-Sanches-Slobin generalization, insofar as number marking is optional, it certainly casts doubt on the received view which sees it as a rare phenomenon. But the co-occurrence of plural and classifier illustrated by Krasnoukhova for Miraña (Boran) and Baure (Arawakan) has a stronger theoretical significance, since number marking is reported to be obligatory in the former language (3.2.1) and obligatory for human and animate nouns in the latter language (Krasnoukhova 2012: 110). Quite

apart from the need to make precise the sense of "optional" in order to test the predictive value of the generalization, these examples, and possibly those in the recent varieties of Nivkh illustrated by Gruzdeva, clearly deserve further investigation.

Classifiers, we may recall, are morphemes which in the relevant languages connect nouns to their morphosyntactic context, in particular to counting structures for numeral classifiers. Whether or not their presence implies that nouns, in the relevant languages, conceptualize their reference as a continuous mass (a conclusion that seems increasingly unwarranted), it is a simple observational fact that classifiers are grammatically necessary in expressions that denote bounded individuals. In a similar way, some languages morphologize an opposition between nouns that, by themselves, denote pluralities or granular substances (like ants or sand), and a singulative variant of them, which denotes instead a single member of that set or an element of that granular substance (like a single ant or a grain of sand). The issue of how number can fit a language that makes systematic use of classifiers, then, is directly related to (if distinct from) the issue of how number can fit systems with a systematic opposition between singulative and non-singulative forms of nouns. Our sample confirms that a number system based on contrasting values like singular, dual, and plural (or others) can indeed coexist with the opposition between nouns that are inherently singular-denoting and nouns that are inherently plural- or mass-denoting. The two values of this opposition are usually referred to as "singulative" and "collective", but Jakobi and Dimmendaal (this volume, on Karko, Nilo-Saharan) use the term "plurative" for the plural contrasting with the singulative, in keeping with Africanist research; notice that Fassi Fehri uses "plurative" in a different sense. Arabic and Karko offer the best illustrations of the sometimes very complex interplay of morphology, morphosyntax, and semantics in such systems; Nivkh and Ket (Yeniseian) also show how the second opposition can coexist with the main singular-plural one, at least for a subpart of the nominal lexicon. It is important to emphasize that the two are indeed distinct, not only in form but also in meaning. While singular and plural differ only in terms of cardinality, singulative and "collective" differ primarily in terms of individuation: the singulative does not simply denote a singular item by opposition to a collection, but an individual unit out of a mass-like (not countable) plurality consisting of uniform items, like a particulate substance (rocks, fish eggs / roe, sand, mosquitoes are among the notions so described); correspondingly, the singulative may be pluralized, as Vajda and Fassi Fehri expressly note for Ket and Arabic respectively. At the same time however, in systems like Karko (common in the Nilo-Saharian family; see Dimmendaal 2000), singulatives simply mark the singular, for a substantial subpart of the nominal lexicon. The two functions of the singulative as individualizing an atom (which may itself be pluralizable) out of a substance-like uncountable plurality, and as making up by derivational means the noun form filling the singular cell in a paradigm, must be sharply distinguished, as emphasized recently by Kouneli (2020); but the very need for this reminder shows how deeply intermingled they can be.

4.2 General number

Various chapters in the volume refer to the use of plural marking as optional rather than obligatory (Vydrin, this volume; Gruzdeva, this volume; Krasnoukhova, this volume; Shimoji, this volume). In these languages, a form of a noun that is formally unmarked and contrasted with a formally marked plural form also seems to be functionally unmarked; in Corbett's (2000: 10) words, such unmarked nouns may be "non-committal" as to the plurality of the referent. With a reference to Andrzejewski (1960), to model this pattern Corbett uses the concept of *general* number, applied to systems where "the meaning of the noun can be expressed without reference to number" (Corbett 2000: 9–10).

Systems involving general number are in contrast with the use of number in e.g. European languages, where the form opposed to the plural seems to be necessarily singular in its reference even if formally unmarked, and are relatively widespread crosslinguistically (see many references in Corbett 2000 as well as *optional number* values in Haspelmath 2013). In this volume, the only chapter that very explicitly states that the grammatical category of number in a language outside Europe is obligatory in the same sense as in European languages is that on Tswana, Southern Bantu (Creissels, this volume); we suspect the explanation might be that there is no nominal form which is formally unmarked for number because number is cumulated with gender which is obligatorily marked on nouns (simplifying slightly), similarly to the way a noun cannot be formally unmarked for number in a language like Russian, where it is cumulated with case. Informally, expression of number in languages such as Tswana or Russian is bound by morphosyntax in the sense that the expression of such obligatory categories as gender (Tswana) or case (Russian) requires overt specification for number; see also Corbett (2000: 12–13) on similar effects from agreement. This is probably not a sufficient condition for obligatory number marking; even in Tswana nouns lower in animacy may be used in their singular form for plural reference. As Creissels (this volume, 2.3.2) puts it, "nouns that stand high in the hierarchy are more prone to be treated as count nouns". It does seem, however, that systems with the singular which is formally unmarked tend (though do not have) to belong to this typological stock, while systems with a marked singular apparently rarely if ever use it as general number. Corbett (2000: 16–18) further notes that the only attestation of the general number formally identical to the plural is in those languages – and in those nouns – which have plural as their unmarked form (cf. Jakobi and Dimmendaal, this volume, on tripartite systems, and Harbour and McKenzie, this volume, on inverse number). With all examples discussed by Corbett, and with all examples of optional number marking in this volume in mind, it remains an empirical challenge to find an example of a language with an overtly marked general number form. But it is the very notion of general number and optional plural marking that we intend to revisit in this section. Note that we leave the discussion of optional duals and trials until later, where we argue that the use

of the term *optional dual* and *trial*, as opposed to *optional plural*, may indeed be appropriate (see section 4.6).

There is a certain ambiguity in the description of such systems of number marking as *optional*, in the sense that this leaves open the question of how many nominal forms should be posited in their (most typical) formal design. Are there only two forms, including one unmarked (the one which Corbett describes as "non-committal") and the other one, marked for plural? Or are there three forms, including one formally unmarked singular, the other marked for plural and the third one, unspecified for number and unmarked for this category altogether (the same "non-committal" form, also described by Corbett as lying "outside the number system"), with the first and the last one being formally identical? While we cautiously – see below – assume that Corbett's general number leans towards the latter interpretation, it is probably the former that is generally implied in language descriptions that refer to the category of number as optional. In this section, we are going to focus on this ambiguity.

To put the concept of general number in a sharper perspective, Corbett (2000: 10–11) uses data from Bayso, a Cushitic language and Fula, an Atlantic-Congo language that distinguish three different forms, singular (formally marked), plural (formally marked) and general (formally unmarked) (see also Corbett and Hayward 1987). He then extends the notion of general number to the languages where number marking is described as optional. These systems are modeled as lacking a dedicated general number form, combining the general number with the singular. Many languages in this volume belong to this type, including some also mentioned by Corbett, Japanese (Japonic) and Arabic (Semitic). Fig. 5 compares the two types of systems featuring general number with the systems which lack it.

unmarked	'general'			'general'			
marked	'singular'		unmarked	'singular'		(un)marked	'singular'
marked	'plural'		marked	'plural'		marked	'plural'
Type a. Bayso, Fula (General number formally different from the overtly marked singular)			Type b. Japanese, Turkish (General number formally merged with the formally unmarked singular)			Type c. Slovenian, English (No general number)	

Fig. 5: Types of number systems with respect to general number (based on Corbett 2000: 9–10).

In Fig. 5 Type a, both singular and plural are marked, but the marking is different; and in Type c, whether the singular is formally marked or unmarked, it specifies singular reference. In other words, by *marked* here we mean formal marking, i.e.

the respective value being expressed overtly, and by *unmarked*, a value which lacks such overt marking.

Apparently, languages belonging to Type (a) feature the category of general number, while languages belonging to Type (c) lack it. As to languages belonging to Type (b), under all readings of Corbett's analysis, two different comparative concepts (or etic categories), general number and singular, are similarly expressed by absence of marking. What is unclear is whether these two different comparative concepts also correspond to two different descriptive (or emic), language-internal categories of general vs. singular number that formally coincide in the unmarked form; or there is one single category which is number-neutral and formally unmarked.

This is not explicitly discussed by Corbett. In his survey, the boundary is not always clear between typological analysis and analysis of individual linguistic systems, or whether the difference is there in the first place. For Turkish (Turkic), he describes the opposition as "general/singular vs. plural (Turkish *ev* 'house'/'houses' vs. *evler* 'houses') in which the first form does not by itself establish a number for a noun", while earlier at the same page, in the case of Japanese, he suggests that "without marking, *inu* 'dog' does not specify for number" as opposed to *inu-tati* 'dogs' (Corbett 2000: 14). The wording is different ("general/singular vs. plural" for Turkish; "does not specify for number" for Japanese), but both languages are intended to exemplify the same type.

Based on the general context of the discussion, we tentatively assume that Corbett's general number does not introduce a single form that, in a language like Japanese, can have either singular or plural interpretation. Instead, a general number form is not interpreted in terms of numerosity at all. It refers to an entity or entities of the kind but is number-neutral in a way nouns are in languages lacking the category of number altogether; it is in this sense that lies "outside the number system" (Corbett 2000: 10). In other words, the combination of the comparative concepts 'general' and 'singular' in one form does not result in one descriptive category of number-neutral form but is a formal merger of two nominal forms whose relation to the category of number is different. As opposed to this view, characterization of number marking in such languages as *optional* is at least compatible with assuming only two distinct forms: one, unmarked both formally and functionally, non-specific with respect to number, the other, marked, specifically plural in its reference.

The model positing three forms may be viewed in the following way: speakers either do not care for quantifying the reference of a noun (then using a number-neutral form) or, if they care, they use either a plural or a singular form of a noun. In the latter case, following the cross-linguistic tendency not to mark singular reference overtly, in most languages the singular form of the noun happens to be identical to the number-neutral form, which is also naturally unmarked. This does not happen in Bayso exactly because Bayso has overt marking dedicated to the singular,

so that the singular form is distinct from the number-neutral form. Fig. 6 compares this view of general number (on the left) with the representation probably tacitly accepted in many cases where such systems are interpreted as optionally marking nouns for plurality (on the right).
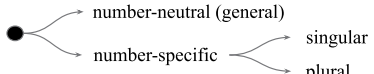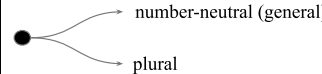
						singular reference	plural reference
number-neutral	unmarked			number-neutral (unmarked)		+	+
number-specific	unmarked	'singular'					
	PL	'plural'		PL		-	+
General number as a distinct descriptive category (number-neutral forms separated from formally unmarked singular forms)				Optional plural marking view (number-neutral forms merged with singular forms into one descriptive category of 'unmarked number')			

Fig. 6: Two ways to model systems with general number.

On the left pane, general number systems are represented as having three forms of a noun, the general number form, the singular form and the plural form. On the right pane, the same pattern is described as only having two forms of a noun, the unmarked one, which allows both singular and plural reading, and the marked plural form, which may only be interpreted as plural. Is there a way to decide which of the models is more descriptively adequate? We can only defend the view that general number and the formally unmarked singular constitute two separate descriptive categories in a specific language in case the distinction is relevant for its grammar. We are next going to argue that this is indeed the case, at least for many languages and at least under some definitions of grammar.

We believe most scholars of number will agree that there is no such thing as optional plural marking in a discourse perspective on the use of grammatical categories, similarly to how there is no free variation in a sociolinguistic perspective on language use. If a language has the category of nominal number at all, it is under some contextual conditions that a noun is number-neutral. Under these conditions, the noun lacks the category of number in the same sense as nouns are *transnumeral* in the languages that lack the category of number altogether. As a simple case, there is a general consensus that in many if not most languages incorporated nouns lose the capacity to inflect for number and become number-neutral (see review of the analyses in Massam 2009). Under different conditions, the same noun is number-specific and thus, whenever formally unmarked, strictly singular,

in the same way as it is in languages like English. No value of the number category is functionally unmarked in the sense of being interpretable as either plural or singular in terms of reference. It is a form of a noun which is ambiguous between number-neutral use and the formally unmarked singular, and the addressee is guided by the same conventionalized discourse practices as the speaker who uses it, to interpret it one way or the other. The term *optional* plural marking masks an important empirical issue. To push Corbett's view of general number systems forward towards a research programme, typology needs to study language-specific conditions under which the expression of numerosity becomes relevant in discourse, and therefore an unmarked form acquires a strictly singular reading.

Here, it is also Corbett who gives us an outlook on a possible range of factors. With respect to general number in Bayso, he notices (Corbett 2000: 10, Note 1) that it is incompatible with modifiers in the NP: If a formally unmarked noun is accompanied by a demonstrative, it can only be interpreted as singular. He hypothesizes that general number negatively correlates with a range of discourse-related factors including definiteness, specificity, referentiality, topicality and first reference to an entity (Corbett 2000: 14). Some of these discourse properties may also account for a crosslinguistically much less frequent phenomenon of general number in pronouns (Corbett 2000: 15, Note 9 on Asheninca (Arawakan); Krasnoukhova, this volume, on Qawasqar (Kawesqar), Gil, this volume, 2.2, or more specifically third person pronouns, as discussed in Corbett 2000: 168, note 30 for Gurr-goni (Maningrida) or West Circassian in Bagirokova et al. (this volume).

As Corbett further indicates, to account for these discourse preferences, we may at some point need to also consider the lexical meaning of nouns. In this volume, Vydrin mentions that number marking in Eastern Dan (Mande) is more frequent with human nouns than with inanimates (see also Gruzdeva, this volume, 2.3.1 on Nivkh, isolate; Creissels, this volume, 2.3.2 on Tswana, Southern Bantu). Apparently, nouns lower on the Animacy Hierarchy may be less prone to show number distinctions because they are less salient in discourse. But the impact of the distinctions in animacy on number marking that initially reflected discourse salience may eventually become lexically categorical, shifting from discourse properties of NPs to inherent properties of their heads. Corbett (2000: 15) mentions that, in Iranian languages, the general number is not observed with human nouns. Languages with "optional" plural marking limited to inanimates are shown as a separate type in Haspelmath (2013), where different cut-off points are also explicitly discussed (major divisions being grouping animate non-humans with humans or with inanimates). In this volume, the typological variation in the availability and obligatoriness of number marking in the languages of South America is discussed by Krasnoukhova. Following Corbett's analysis of the data from Iranian, the category of general number in these languages is understood as limited to a certain lower segment of the Animacy Hierarchy, and different mapping of obligatory number marking on the top of the Animacy Hierarchy represents different patterns of gram-

maticalization of the tendency to constrain nouns typically salient in discourse to number-specific use.

To sum up, the conventional notion of *obligatoriness* of number marking high on the Animacy Hierarchy as compared to its *optionality* lower on it and to *absence* of number marking yet further down may be viewed as conventionalization of discourse preferences that, in some languages, may rule out number-neutral use for nouns of certain semantic classes and number-specific use of nouns of certain other classes, turning these preferences from grayscale "maybe" in Eastern Dan to black and white "yes or no". Cf. the following Fig. 7 where several possible patterns are shown as examples.

	human	animate	inanimate
a) Eastern Dan (Vydrin, this volume)	number-specific	↔	number-neutral
b) Iranian (Smirnova 1981, via Corbett 2000: 15)	number-specific only	number specific ↔	number-neutral
c) Trumai (Guirardello 1999, via Krasnoukhova, this volume)	number-specific ↔	number-neutral	number-neutral only

Fig. 7: From discourse preferences to "obligatorification" of number marking along the Animacy Hierarchy.

The conventional characterization of plural marking as *obligatory* corresponds in Fig. 7 to the segments of the Animacy Hierarchy where only number-specific uses are available, lack of nominal plurality corresponds to the segments where only number-neutral uses are available, and *optional* plural marking corresponds to where the two overlap and compete, with preferences distributed over morphosyntactic and discourse conditions as suggested by Corbett and likely to increase along the Animacy Hierarchy towards more number-specific uses. It is in these areas of overlap that language-specific grammars are sensitive to the difference between number-neutral unmarked general number form and formally unmarked singular form. This is clear in the case of morphosyntactic conditioning, because failure to take it into account leads to ungrammatical sentences (as when using an unmarked NP with a demonstrative referring to several entities in Turkish). But it is also arguably so in the case where also or only discourse conditions are at play; indeed, in terms of efficiency of communication, both the speaker and the addressee must take them into account when producing and perceiving an utterance, otherwise an NP intended as "non-committal" may be interpreted as singular, or vice versa. Thus, the decision whether two or one unmarked form need be posited in a language may ultimately depend on whether we consider discourse organization to be part of language grammar.

As one candidate case for morphosyntactic conditioning of general number, consider numeral phrases. In this volume, plural marking is often reported to be ungrammatical or infrequent with numerals, as in Nivkh, Nganasan (Uralic) and Kakataibo (Panoan), or with lower numerals, as in Ket (Yeniseian). This salient typological parameter is discussed by Krasnoukhova (this volume) for languages of South America and Velupillai (this volume) for contact languages, considered by Matasović (2000) in a wide cross-linguistic perspective and briefly mentioned in Plank (1989). Some contributions specifically discuss factors of variation (Vydrin, this volume, 3.4 on Eastern Dan; Bagirokova et al., this volume, examples (95)–(98). In such contexts, the lack of the plural marking on a noun is best viewed as a number-neutral form. Importantly, the number-neutral form may be restricted to numeral phrases. Thus, East Caucasian languages are not usually described as languages with general number, because the noun formally unmarked for plural implies singular reference. Yet, the noun in numeral phrases is in the unmarked "singular" form; or, more precisely, it is most often in the "singular", with the exact conditions of plural marking poorly explored (see Haspelmath 1993: 232 for a brief discussion of Lezgian, East Caucasian).

While this line of argument is fully compatible with Corbett's account of general number, it also provides a possible reason why his discussion of the phenomenon is not particularly clear as to the choice between the two schemes in Fig. 6, left vs. right pane. He explicitly says that: "(W)e have to consider the form–meaning pairings available [to the lexical item – added by the present authors], independently of context." (Corbett 2000: 14). It is clear from the discussion above, including much of Corbett's own evidence, that it is not only the lexical item itself but also the context that controls the choice between number-specific and number-neutral uses of the noun, and consequently singular or non-specific reading of a formally unmarked noun. As long as we only consider inflectional capacity of a lexical item, we cannot prefer the interpretation of systems with the general number as involving two unmarked forms over their interpretation as involving only one such form. In terms of morphological inflection, such systems are conveniently described as only having the total of two forms, one ambiguous as to plurality, the other confined to plural reference. In purely morphological terms, the expression of number is indeed optional (see also Shimoji, this volume, on number in Japonic languages).

Terminologically, Corbett's *general number* might not be the best term because it suggests understanding the phenomenon as a special value of the grammatical category of number, and is easily interpreted as describing a form of a noun which can be interpreted as having either plural or singular reference, thus shifting towards the notion of optional number. Fig. 6 (left) and Fig. 7 put into relief the logic of only a formal 'merger' of the singular with the general number form.

To sum up, in this section we have argued in favor of making a distinction, at the level of language-specific categories in languages conventionally described as featuring *optional plural marking* or *general number*, of two identical forms, one

number-neutral and the other singular, both formally unmarked. Note that, within the formal framework, this is the solution Bylinina and Podobryaev (2020) independently arrive at in their analysis of general number in Buriat (Mongolic). We suggest that the unmarked singular form is paired with the marked plural form, making this opposition comparable to that of the languages with "obligatory" number marking. The unmarked number-neutral form is comparable to the noun in languages that lack the category of number altogether. We argued that grammars of such languages are sensitive to the distinction of the two unmarked forms of the noun, because in some contexts a formally unmarked form is unambiguously interpreted as singular in reference, which reflects language-specific rules. The two forms are distributed across discourse conditions and sometimes also morphosyntactic contexts that all can be roughly described as related to prominence in discourse. It is from these rules that categorical lexical rules of availability of number marking probably arise.

This revision, or maybe more modestly recapitulation, of Corbett's notion of general number, together with an explicit parallel with transnumeral nouns in languages lacking grammatical number, helps us focus on the factors of the variation in the expression of number instead of essentially dismissing them by describing number marking simply as optional, and profiles a specific empirical question, which is: Under which discourse / pragmatic conditions may a NP be non-committal as to the number of its referents? Answering this question would require a battery of corpus or experimental techniques and, while prompted by the discussion of optional number in the contributions to the volume, definitely lies outside their scope, so we leave this track open to further investigations, including for the languages in the volume.

4.3 Cardinality-based values

Several chapters evidence a particular status for the paucal, and generally for values other than singular and plural, or varieties of these that are more precisely qualifiable in terms of 'minimal' and 'augmented' (see Harbour 2011, 2014).

The clearest evidence of a qualitative difference between number values involves duals, trials, and quadrals. In abstract terms, these are simply specifications of cardinality, not different from the singular in this respect; but even a basic acquaintance with linguistic facts shows that number values are not just restrictions on the cardinality of the denotation. For a system of numerals, any cardinality is equivalent to any other for grammatical purposes, maybe with some extra determinations for the cognitively more salient values of one and other "low" units; for instance, in German (Germanic branch of Indo-European) only the cardinal numerals 1, 2, 3 are inflected for case. Number systems, on the contrary, primarily oppose being one to being more than one. Besides, the singular does not necessarily corre-

late with a count denotation whose cardinality is 'one', as nouns like *air* or *tiredness* show. In addition, only the reference to cardinality 'two' has a widespread linguistic reflex in a dedicated dual number value. We will reconsider below in section 7 the range of interpretations associated with singular and plural, and we will see in section 4.6 that denoting two-membered sets cannot be the defining property of the dual, since that value systematically expresses more than a mere cardinality restriction. As for values corresponding to other cardinalities, it is amply known that these may only extend to trials and possibly quadrals (on controversial quadrals, see Corbett 2000: 26–30), in very few systems clustered in the Pacific. Among the languages included in our sample, only Solomon Islands Pijin (Velupillai, this volume, 2.2.2) and Russian Sign Language (Kimmelman et al., this volume, 2.2) evidence a trial, in the latter case also a quadral, both restricted to pronouns and both transparently incorporating a numeral. The scarcity of such values and its restricted expression are entirely expected, as Krasnoukhova notes (this volume, 2.2; they are known to be largely limited to the Pacific area). These are ways to integrate a numerical characterization into referring expressions ("two – them"), not separate number values.

4.4 Paucals and "approximative size" values

Many number-inflected nominals characterize the size of their denotation in general terms, as small, large, or extremely large pluralities, but not all of them are necessarily instances of specific values corresponding to these interpretations. The 'small plural' or 'large plural' readings available for Arabic are probably the most straightforward example of this arising from our sample. As Fassi Fehri (this volume) argues for Arabic (Semitic), these are effectively a by-product of the open-ended nature of plural formation processes in that language. Mass nouns can often be pluralized, and the plural reading that retains a mass interpretation (rather than being recategorized as a count noun denoting a type) is one of abundance. For count nouns, a double pluralization likewise leads to a 'larger plural' reading, as in *xuruuq-aat-un* '[many] violations', where *xuruuq* is a "broken" plural, derived from the singular *xarq-* by a rearrangement of the CV pattern, which is in turn suffixed by the plural *-aat* (Fassi Fehri, this volume, example (24)). But this does not amount to recognizing a distinct subcategory of abundance plural as a *grammatical* category, as this type of readings do not for example determine a separate agreement pattern. In keeping with this interpretative standpoint, Fassi Fehri (section 2.3) expresses caution about the traditionally reported categories of "plural of paucity" and "of abundance", noting that they were more likely appropriate as labels for the content of alternative broken plural forms rather than subvalues of the plural number value.

Marori illustrates another way in which a system may express a reading that we can justifiably call 'paucal' without this being one of the available values. The reading arises as a result of the interaction between number marking on noun phrases

and on verbs (which, like in several other cases noted in this volume, is not a case of agreement but of two independent and indeed different loci for expressing the number of participants denoted by the verb's arguments). In particular, when a NP marked as plural expresses the argument of a verb that indexes it by means of a singular affix, the NP is interpreted as denoting a small plurality (Arka and Dalrymple, this volume, example (50)). Nowhere else in the language does a paucal exponent minimally contrast with other number values. Indeed, in the expression of verbal plurality, Dalrymple and Arka note explicitly that the value 'non-plural' does not correspond to a denotation of sets with cardinality one or two, as for the nominals, but more loosely expresses a 'small number' (section 2.4).

The paucal of Lower Sepik languages (Foley, this volume), on the other hand, is to all appearances a value alternative to singular, plural, and dual: some languages in this family have a dedicated expression for it on pronouns, nouns, and verbs. However, even in this context, the paucal is set apart from other values. Lexical nouns in Chambri express a paucal reading by adding the suffix *-reb* to an already pluralized noun (see his example (12), section 2.2). The impression of paucal as a morphologized specification of plural is reinforced when we note that in Yimas even a noun marked as plural triggers a paucal interpretation when in construction with a verb that carries itself the generalized paucal suffix (again 2.2; notice the parallel with the Marori case just discussed). This is not to question the morphological evidence for a distinct paucal marking: but even in this case, a blunt statement to the effect that these languages encode a singular, dual, paucal, and plural would be simplistic. It would be correct, insofar as the language makes these distinctions by grammatical means; but it would conceal the fact that paucal and plural can coexist syntagmatically, without contrasting the way singular, dual, or plural do with each other.

The phenomena we have considered lead to a clear but nuanced conclusion. The paucal is definitely a possible number value, as shown, for instance, by the fact that it can qualify nouns in Gurindji Kriol, and it is one of the ways verbs specify the numerosity of their arguments in Gooniyandi (see respectively Velupillai, this volume, and McGregor, this volume). However, its use should be documented carefully, to clarify whether it is used systematically, for collections of the appropriate size, as an alternative to singular, dual, and plural, or whether it is more like an extra characterization that qualifies the plural as a core number value.

4.5 Collective and distributive

The various mentions of special collective and distributive forms motivate a different conclusion. We are not talking here about group-nouns or particular readings of plurals, but of dedicated morphological marking for such readings. Here the evidence shows that even when a language marks such readings morphologically on

nouns, these forms usually express a property of the situation, which is distributed over distinct dimensions (times, participants, places), but it is not one value of a paradigmatic set of oppositions comprising singular and non-distributive plural. A formal means with which the distributive reading of a plurality seems to be strongly associated seems to be reduplication, as in Indonesian, Malayo-Polynesian (Gil, this volume, 2.4.2), where, along with other readings such as intensification, reduplication expresses distributivity. Kakataibo (Panoan) also uses reduplication but specifically for iterative readings of verbs, thus distribution over times (Zariquiey, this volume, 2.3.3), while a verbal distributive affix "indicates that the action is being carried out by various individuals independently" (2.3.2.1.2), and a NP distributive enclitic "indicates that the event is associated independently with each of the individuals referred to by the NP" (2.2.2.2). Interestingly, a distinct collective morpheme is used only with nouns (denoting animal species or dead ancestors) already marked by the generic plural enclitic, showing that the two characterizations complement each other and are not alternative ways to carve up the paradigmatic space (cf. Mithun 1999: 91 on Yana, Yana). A similar or more restricted function is reported for Mohawk (Iroquoian) verbal or nominal distributive affixes (Mithun, this volume, 2.4, 2.3.1), where the latter specifically denote a multiplicity of types and not of token portions (see example (30)); and for Nivkh, isolate (Gruzdeva, this volume, 2.4, 2.3.3), where reduplication applies to verbs or to nouns, and the latter for instrument participants and forming derived adverbs. The closest the distributive / collective characterization comes to a systematic morphological alternation is probably in the pronouns of Russian Sign Language (Kimmelman et al., this volume, 2.2), which are reported to be generally characterizable as collective or distributive (and here too, reduplication may apply to nouns and verbs, see section 2.3.3).

These facts lead us to view distributive or collective marking as grammatical characterizations that complement those expressed by "core" number values. The paucal is also somewhat peripheral, in that it expresses numerosity as a qualitative property of pluralities as being "small", which may entail that when this marker is absent the plural is implied to denote a "large" plurality, as Foley reports for Yimas (this volume, 548). The dual is where this qualitative type of characterization and the quantitative one intersect.

4.6 The dual

Our sample shows that the dual is anything but a rarity, since it is absent as a dedicated value of number only in West Circassian (Northwest Caucasian), Karko (Nilo-Saharan), and Tswana (Southern Bantu); note that in Nivkh (isolate) and Eastern Dan (Mande) the dual is only represented as the dual inclusive pronoun, a category possibly independent of the dual as linguistic expression of cardinality two; see section 3 above). Occitan (Romance), which lacks a morphological dual, dis-

plays a "quasi-dual" form of a plural indefinite determiner referring to a single bi-partite object like a pair of pliers (Bach, this volume, 2.3, examples (2)–(3)). When forms interpretable as paucal are attested, they are distinct from duals. In short, the legitimacy of the dual as a distinct value of number, denoting sets of cardinality two as opposed to both singularities and pluralities (including small pluralities), is not in question. What calls for comment is the fact that this value does not correlate in a straightforward way with denoting sets with cardinality two. The dual is known to be often restricted in its distribution, not only categorially (not being equally available to pronouns, nouns, verbs) but also lexically (limited to some semantically determined subareas of the lexicon), insofar as it can presuppose a cohesive inter-pretation (*ambal* dual in Plank 1989, see also Corbett 2000: 20), and to be suscepti-ble of uses where it denotes small but not two-membered cohesive collections (such as 'teeth', expressed by a formally dual form in Akkadian (Von Soden 1969) and in Modern Hebrew (Schwarzwald 1991), both Semitic.

The accounts in the contributions to this volume confirm this characterization. Just like the other number values, when the dual is a value for pronouns, it may be unavailable or very restricted among nouns, not just where nouns are generally unmarked for number, like in Indonesian (Malayo-Polynesian) or in Marori (Trans-New Guinea), but also elsewhere. For example, it is unavailable for nouns in Kaka-taibo (Panoan) and it is available but limited in Japonic languages and Ket (Yenisei-an); Southern American languages and contact languages confirm the same pattern.

Even where the dual is available, closer inspection may reveal non-trivial limita-tions in its distribution. In Arabic (Semitic) pronouns, only second- and third-person pronouns have a distinct dual form, but not the first (which is notable, since the Animacy Hierarchy would rather lead us to expect the first and second persons to display most number oppositions); in Kakataibo, only older speakers retain a "dual" form of pronouns which is interpreted as dual for second person and as paucal for third person; in Marori, pronominal suffixes on the verb stem display a dedicated dual form only for the first person (this is limited to Actor arguments), while for the second a dual interpretation is constructed by the simultaneous presence of non-singular and non-plural affixes (*constructed* number values in Corbett 2000: 169); and in Kiowa (Kiowa-Tanoan), inverse marking can express a singular or plural but not a dual value.

Significantly, several authors remark that the plural may occur instead of the dual to denote two-membered collections (Arabic, Papuan Malay and Kakataibo; see Corbett 2000: 2.3.3, also on optional trials), thus making the dual an *optional* value. In Nganasan (Uralic), in numeral phrases with 'two', the noun is usually unmarked, but can also be marked as non-singular. In this latter case, it may take either dual or plural marking. Duals that alternate with plurals, which we consider truly optional, are different from the concept of "optional" plural marking discussed (and dismissed) in section 4.2 above. The optional dual is probably best viewed as a possibility of morphological specification of the exact cardinality for non-singular

referents, similarly to how cardinality may or may not be lexically specified by numerals, including by the numeral 'two' in languages that lack a morphological dual. Such lexical specification of cardinality two is optional by definition – otherwise the relevant item would probably qualify as a *dual word*, a subtype of *plural words* (Dryer 1989). Here, we want to emphasize the difference between the core number opposition of singular vs. plural and additional cardinal number values such as dual or trial, which are conceptually close to numerals. This affinity is reflected in the existence of optional cardinal values that may be viewed as intermediate between "pure" grammatical number and lexical means of cardinality specification.

Most of these distributional patterns – preference for pronouns, optionality – are compatible with a view of the dual as just a value of the category of number on a par with its other values. What really makes dual special is its content (Plank 1989, 2.5). A very clear case is provided by Gooniyandi (McGregor, this volume), where the dual NP enclitic, like its plural counterpart, presupposes that the pair (or plurality, for the plural enclitic) referred to is not just a collection, but a collection that forms a group ("they [dual and plural enclitics] indicate as well as number also the sense that the referents belong together as a unified group", 3.1). In other cases, a privileged link connecting specifically the dual and a collective / group interpretation emerges from lexical restriction: in Ket (Yeniseian), nouns denoting naturally paired objects contrast a plural denoting a functional pair (so, a collective dual) and a morphologically more complex plural for collections of greater cardinality (Vajda, this volume, 2.3.2).

The interaction between dual morphology and interpretation as a functionally cohesive (collective) pair can also emerge more indirectly, if just as compellingly. Slovenian and Nganasan illustrate languages where the use of the dual is widespread with lexical nouns (as well as on pronouns), *except* for nouns denoting naturally occurring pairs, that is for dyads, especially bipartite body parts (Gusev and Wagner-Nagy, this volume, 2.3.2.1 and Stegovec, this volume, 2.3.3), which instead occur (or, in Nganasan, tend to occur) in the plural. Stegovec analyses in some depth this property of Slovenian, which is all the more intriguing given that such nouns do have a dual form, as for two randomly chosen legs of an insect in his example (99), which typically occurs after the numeral 'two' (the plural is also admissible, but with a special form of the numeral). His conclusion is that the pair nouns do not take a dual form when denoting cohesive pairs because a restriction to two-membered sets is already available as part of their lexical meaning, and so they are semantically "lexical duals", in a sense. In order to express the cohesive reading of "natural body part" (a dyad), they simply use the plural, for which a collective reading is always available, while the "pair" component is encapsulated by the lexical item itself. Again, the dual behaves as an optional category: no need to express cardinality two when it is the default inherent to the noun. Other nouns instead require the explicit dual marking to license an interpretation as functionally related pair, or dyad, as we will call such cohesive pairs.

In addition, Stegovec reports that a morphological dual must refer to an entity already introduced in the discourse: "outside pair nouns, bare dual nouns cannot be used without an already established referent." (Stegovec, this volume, 4.4). This additional circumstance makes the parallel with Nganasan even more revealing, since in that language, too, duals are subject to a familiarity constraint. In this case, the use of the bare dual contrasts with numeral modification by 'two' (which governs the singular): "if the objects in question are mentioned for the first time, they are used with the numeral; when they are already known, dual or plural is enough" (Gusev and Wagner-Nagy, this volume, 3).

Vaillant (1958: 20) observes for Old Church Slavonic (Indo-European) that in this language, too, first reference to a pair tends to include, in addition to the morphological dual, an explicit numeral. This is, however, not applied to dyads, which, unlike Nganasan, require dual marking. Dyads lack a lexical numeral already at first mention. Thus, for dyads, just as the morphological specification of the dual is considered redundant in Slovenian and Nganasan, lexical (but not morphological!) specification of the same cardinality is considered redundant in Old Church Slavonic. This is another cross-linguistic parallel between lexical and morphological expression of exact cardinality two.

The connection between morphological expression of the dual and this familiarity restriction may be interpreted as follows: when dual-marked nominals must refer to familiar discourse referents, this is because the dual has a presuppositional value as compared to lexical expression of the cardinality two. Duality must be asserted when the discourse referent is introduced, and for that, an explicit numeral modification is needed (first mention, foregrounded cardinality, use of the numeral); on successive mentions, this property is part of the shared conversational background and can be referred back to by the use of the dual (further mentions, backgrounded cardinality, use of the dual). However, there can be two sides to duality: the mere property of having cardinality two, and binarity as the property of making up a cohesive pair, where one member presupposes the other. Stegovec shows that Slovenian pair nouns are plural when they denote natural pairs, but they can be dual when they denote two-membered sets that do not make up a natural pair (again, two randomly chosen legs of an insect, in his example (99)); he hypothesizes that they 'assert duality as a matter of their lexical semantics', and so do not require duality to have been introduced in the discourse, nor do they require a dual marking to express the property of having cardinality two. The cohesive, natural-pair reading inherent in the lexical predicate can be suspended, however (in contexts like that above, where two legs may not be a natural pair), and that is when the dual is allowed. Leaving aside specific language-particular details, what matters is that in languages like Slovenian and Nganasan dual marking signals a presupposed reading as 'having cardinality two', and it can apply generally without bringing about a cohesive interpretation for the nominal it marks; however, it is sensitive to such a reading if it is already lexically encoded. And when we note that classical Indo-

European philology had already recognized both the cohesive reading and the anaphoric nature of the dual (Brugmann 1900: 371, Vaillant 1958: 20), we can confidently conclude that this interaction of semantic properties is one of the ways the dual manifests itself in language, and not a fortuitous language-specific occurrence.

Conceptually, cohesion is distinct from cardinality two. But natural cohesion almost exclusively applies to cardinality two, because some real-world entities come in pairs but very few to none in triplets etc. As a result, natural cohesion grammatically interacts with the cardinality two more than with any other cardinality. We follow Plank (1989) – who declared he followed Humboldt (1827) and Jespersen (1924) – in distinguishing types of dual on the basis of their interaction with cohesion. Some languages limit the use of the dual to dyads (type Gooniyandi), some languages limit the use of the dual to cardinalities two *except* dyads (type Nganasan or its distant relative, another Uralic language Khanty, Nikolaeva 1999: 12–13), and some apply the dual to all cardinalities two independently from how cohesive they are (type Old Church Slavonic; see Suprun and Moldovan 2005: 49). Further examples can be found in Plank (1989).

	Cohesive	Non-Cohesive
Gooniyandi (Bunuban)	ambal	('two')
Old Church Slavonic (Slavic)	dual	
Nganasan (Uralic)	(plural)	dual
Khanty (Uralic)	(singular)	dual

Fig. 8: Scope of the dual over cardinalities two.

As Fig. 8 shows, Old Church Slavonic has 'pure' cardinality two as the grammatical meaning of the dual and is insensitive to cohesion, and thus is distinct from Gooniyandi where the grammatical meaning is indisputably that of a cohesive dyad, or ambal. On the other hand, notwithstanding their apparent distribution, we are going to argue that the categories of dual in Khanty and in Nganasan are not the opposite of the cohesive dual of Gooniyandi, i.e. not dedicated non-cohesive duals, even if both languages are sensitive to cohesion, each in its own way.

Indeed, Khanty may not use the dual of nouns which refer to dyads because it conceptualizes dyads as units, including referring to elements of such units as halves. Another, and a much closer, relative of Nganasan, a Uralic language Nenets, has a dedicated lexical item to refer to 'one of a pair' (Tereščenko 1965: 37). This is clearly not a special property of their dual categories, because a similar strategy of referring to dyads by singular forms and to elements of such dyads as halves is also reported for Nivkh (isolate), a language which lacks the dual altogether (Gruzdeva, this volume, examples (46)–(51), though the plural is also possible, probably implying ex-

haustivity, example (48)). In Khanty and Nenets, dyads fall outside the functional scope of the dual not because the dual is specifically non-cohesive but simply because dyads are expressed by a dedicated *unitary* strategy. By way of comparison, the plural in languages that also have a dual and a trial is, for typological purposes, still a plural, not a special more-than-three category (see the Number Hierarchy in Corbett: 2000, 2.3). As to Nganasan (and Slovenian), as argued above, it may not use the dual of nouns which are inherently dual because this is redundant, the meaning of cardinality two being inherent to the lexical meaning. This constraint, again, does not come from a semantic specification of its dual as non-cohesive but from the avoidance of double indication of the cardinality two. As opposed to these two languages, in Old Church Slavonic the cohesion is simply irrelevant, and the dual must be applied to any sets of two, cohesive or not. In short, we argue that non-cohesive duals are the unmarked instantiation of the dual as a cross-linguistic category, whose typological variation is thus better viewed not as cohesive vs. non-cohesive but as cohesive, or ambal, vs. unmarked dual (i.e. cardinality of two further unspecified).

The distinctions related to cohesion, however important, are probably not categorical and not diachronically stable. Both Khanty and Nganasan may also use dual with such nouns, though at least in Nganasan this use is marginal. On the other hand, in Old Church Slavonic, it was with the dyads that morphological forms going back to the dual persisted longest, long after the dual as a functional category opposed to the plural was lost in most Slavic languages (Meillet 1934: 369; see also discussion of the plural marking of dyads in Ket earlier in this section).

To sum up, the analyses in our sample allow us not only to review the availability of the dual across several quite different systems; they also allow us to bring into better focus what makes the dual a particular number value. Denoting sets of cardinality two is certainly the core content, but this may be accompanied by characterizations that have to do with the cohesion of the denoted dyads, and may restrict the denotation to contextually familiar discourse referents. Two is not a cardinality like any other because many two-membered sets (but not many higher-cardinality sets) are conceptualized as a bipartite unit, and also because a pair is the smallest and the most fundamental type of plurality, which is already implied by the very distinction between speaker and addressee (a point already recognized in von Humboldt 1827; cf. Plank 1989: 311). It then stands to reason that a single sub-element of such cohesive pairs, conceptualized as bipartite units, may be referred to as 'half', not only in Khanty, but, in our sample, also in Ket and Nivkh (and, incidentally, in Irish). With the dual, a language can grammatically express not so much the concept of being two in number, as the concept of being "twin", and the interplay between the two sides of the dual may give rise to different grammatical patterns.

5 Pronominal and nominal number: a comparison

Throughout the volume, and elsewhere in this conclusion, multiple references have been made to the Animacy Hierarchy as a major principle underlying number marking in individual languages. Roughly, the higher on the hierarchy, the more expected are number distinctions, and the more likely they are to be obligatory, if present. The importance of the differences along the dimension of animacy has been suggested in Smith-Stark (1974). The strongest and most influential proponent of the hierarchy in the domain of the typology of number is probably Gil's *referential hierarchy* (this volume); Greville Corbett (2000: 54–88, 2010); see also Dixon's *nominal hierarchy* (1994: 85), Gil's *referential hierarchy* (this volume), and Croft's *extended animacy hierarchy* (2003: 128–132) where *extended* accounts for the fact that not all differences can be appropriately described in terms of animacy, as first vs. second person or human nouns vs. non-human animate nouns. An extensive cross-linguistic survey of splits is carried out by Ivani (2017, 6.3 and 6.5), see also the database (Ivani and Zakharko 2019).

Vast evidence supports the relevance of the animacy dimension for the occurrence of nominal plurality, with nouns higher in animacy (more similar to humans) showing plurality distinctions, or showing them obligatorily, and nouns lower in animacy lacking means of pluralization, or showing "optional" number marking (Corbett 2000, Haspelmath 2013). Here, we leave aside the interpretational issue whether this is motivated by the fact that human reference is more important to cognition and discourse, or by the fact that human referents are more easily individuated, or both; see section 4.2 for some discussion. Note that Smith-Stark himself suggests that the property underlying plurality splits is not animacy per se but "the likelihood of participation in the speech event" (Smith-Stark 1974: 657). We take the variation along the hierarchy as an empirical point, also substantiated by the data in the volume. In response to section 2.3 of the Questionnaire, the authors of the volume mention the Animacy Hierarchy as a factor in Tswana (Southern Bantu), Eastern Dan (Mande), Lower Sepik languages, Nivkh (isolate), Gooniyandi (Bunuban), in terms of frequency of nominal plural marking.

It is also uncontroversial that number in pronouns is visibly more crosslinguistically common (nearly universal) than number on nouns (common but by far not universal). Comparing WALS data on pronominal (Daniel 2013) vs. nominal (Haspelmath 2013) plurality suggests that there are quite a few languages that distinguish number in pronouns but not on the nouns, but we have no evidence for languages that show number distinctions in nouns but not in pronouns. In the languages of the volume, number is obligatory in pronouns but severely restricted on nouns in Marori, Trans-New Guinea. Varieties of Indonesian differ in this respect, with number being obligatorily specified only in the first person in Jakarta Indonesian, in the first and second person in Riau Indonesian, in the first, second and third person animate only in Papuan Malay, and for all pronouns but not nouns in Kupang Malay (Gil, this volume, 2.2). See also the discussion of the varying scope of obligatory

number specification in Krasnoukhova (this volume). Another common claim is that pronouns make not less number distinctions than nouns: dual on nouns usually implies dual on pronouns, and trial seems to be only attested on pronouns.

We thus have two typological generalizations that seem uncontroversial and robust: (a) pronouns are more prone to number distinctions than nouns; and (b) nouns that are higher in animacy are more prone to number distinctions than nouns lower in animacy. These observations can be further generalized as a conclusion that the category of number is organized along a hierarchy integrating pronouns and nouns, schematically shown in Fig. 9, whose design replicates that in Corbett (2000: 56). Note that there is a range of versions of the Animacy Hierarchy, such as integrating a distinction between higher and lower animates and discrete and non-discrete (mass) inanimates, as in Haspelmath (2013), or adding proper names as an additional cut-off point between third person pronouns and common human nouns as in Croft (2003: 130).

| speaker | addressee | 3rd person | kin | human | animate | inanimate |

preference for number-specific uses
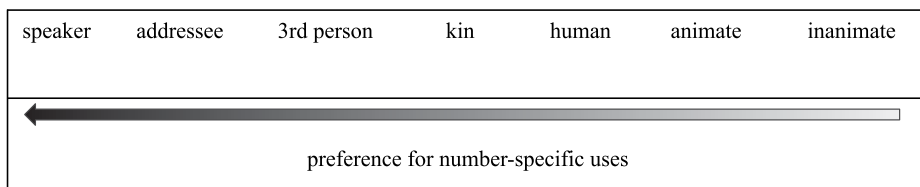
Fig. 9: (Pro)nominal plurality: a unified view on the Animacy Hierarchy.

In Fig. 9, what we mean by number-specific uses is the cross-linguistic occurrence of the category of number on certain semantic classes of nouns and, eventually, its obligatoriness rather than optionality with them (section 4.2). We will not use the hierarchy as a principle predicting the continuity of the areas covered by individual markers in a language, such as cross-linguistic unlikelihood of a marker that would pluralize nouns with human and inanimate reference while another marker is used to pluralize non-human nouns with animate reference. Claims such as that may hold, but unlike the occurrence of nominal plurality, this generalization of formal continuity of the scope of plural markers can be explained merely by the conceptual similarity of the pluralities formed by the referents in adjacent classes. That is, groups of people are more similar to prides of lions and maybe flocks of birds and even packs of fish than to piles of stones or pools of water. While suggesting a conceptual space of adjacency, this kind of similarity does not imply a *hierarchical* ordering of the relevant semantic classes throughout this conceptual space.

Now, what is the evidence in favor of viewing pronouns as a segment on the same hierarchy as the nouns, as they are shown in Fig. 9? Consider the linguistic expression of the numerosity of events (see section 6). Event number belongs to the domain of linguistic quantification together with (pro)nominal number. Yet the

expression of numerosity of events is only rarely formally similar to (pro)nominal number in individual languages (though see reduplication, e.g. in Indonesian; or the use of nominal plural markers on the verb discussed at the end of section 6). The two categories are not fully independent – thus, plurality of participants may imply plurality of events – but still conceptually very different. Event number can be compared to (pro)nominal number along various dimensions, but they are probably never fitted into one conceptual space.

The situation with pronominal and nominal number is not completely unlike this. The two categories do not generally share the same exponence (Daniel 2013). Moreover, plurality in personal pronouns is often expressed lexically by stem suppletion (see Corbett 2005 for a theoretical discussion of suppletion in personal pronouns), while this is much less common with nouns. It is true that in Marori, where only a few nouns have plurals, these are suppletive (Arka and Dalrymple, this volume, 2.3) and thus resemble pronominal plurals. Nevertheless, all other cases of (full) suppletion in nominal number attested in the volume involve isolated lexical items against the background of a more or less regular morphology (Stegovec, this volume, 2.3.2; Kimmelman et al., this volume, 2.3.1, Vydrin, this volume, 2.3.2). Remember that the expression of verbal number by stem suppletion is one of the strong arguments in favor of seeing this as a category independent from nominal number (Durie 1986, Mithun 1988). On the other hand, pronominal number is a means to express multiplicity of entities and is conceptually closer to nominal number than verbal number is to nominal number, at least when the latter expresses numerosity of events. The question is, are pronominal plurals and nominal plurals similar enough in the conceptual dimension and in the dimension of formal expression to be viewed as instantiation of one linguistic category – enough to place them on the same typological scale, the Animacy Hierarchy? In other words, is there empirical evidence or conceptual arguments in favor of considering pronouns as constituting the higher segment of the same hierarchy as nouns, in terms of linguistic categorization of numerosity?

We may start with several points related to the pronominal segment of the Animacy Hierarchy in Fig. 9. First, the hierarchical ordering of the speaker and the addressee as Speaker > Addressee is not universal, as evidenced by the variation across systems of hierarchical alignment (Zúñiga 2006), with one of the best known systems of this kind, the direct-inverse system of Algonquian languages, traditionally described as Addressee > Speaker (Bloomfield 1957, Dixon 1994: 90, Daniel 2005, Daniel 2020b and Cysouw 2005 for a critical assessment). More generally, the mere presence of the inclusive pronoun can be interpreted as evidence in favor of Addressee = Speaker (Plank 1985). Number-related asymmetries between the first and the second person are indeed observed, and evidence is found in favor of a higher preference of number-specific forms in the first than in the second person, as in Sabanê (Nambiquaran), or other asymmetries, such as first person plural being suppletive while second and third person is not in Mosetén, Moseten-Chimane, and

Kulina, Arawan (all three examples come from Krasnoukhova, this volume, 2.2.1) and in Mobilian Jargon and Yimas-Arafundi Pidgin (Velupillai, this volume, 2.2.1). But the evidence for any such asymmetries must be relativized to the person hierarchy observed in the language, a typological study that, to the best of our knowledge, has never been carried out.

Second, even in the small language sample presented in this volume, number distinctions do not always align with any version of person hierarchy and should probably be better considered as idiosyncratic results of a language-specific diachronic development. Thus, in Arabic (Semitic) the dual is attested in the second but not the first person. This cannot be accounted for by Addressee > Speaker hierarchy because the third person, placed below speech act participants in all models of pronominal reference and in all versions of the Animacy Hierarchy, also has a dedicated dual form (Fassi Fehri, this volume, 2.2); see several more examples of the dual absent in the first person in Plank (1989), and Arka and Dalrymple (this volume, 2.4) on number distinctions in the Marori verb. Similarly, the most widely known asymmetry of this kind, English *I* (sg) vs. *we* (pl) in the first person as opposed to *you* (sg&pl) in the second person, which does match the conventional version of the person hierarchy (Speaker > Addressee), is better viewed as a result of diachronic evolution (politeness shifts) than reflecting the Animacy Hierarchy. Indeed, in English, number distinctions are obligatory a long way down the scale in Fig. 9, so that the merger in the second person stands out as an exception. Any evidence for differences in number categorization between the first and second person should be considered with caution.

Whatever the relative ordering of the first and second person, the asymmetries between the first and the second person are by and large irrelevant for the typology of number. As Siewierska notes, "the existence of actual number distinctions only in the first person emerges as being quite uncommon", adding that "[B]y contrast, paradigms in which there is a number distinction in the first and second persons but not the third are widely attested" (Siewierska 2004: 93). She further extends this observation to the number of values, suggesting that, unless one treats the minimal inclusive category as a first person dual (see section 3 above), the most common pattern seems to be that duals, trials and paucals are exhibited by all persons (Siewierska 2004: 95).

Third, in terms of homo- vs. heterogeneity (see section 3) third person pronouns can be argued to belong with nouns rather than with first and second person pronouns. Indeed, their plural is usually viewed as additive and thus different from that in the first and second person (but see Daniel 2020a for a brief discussion of a possibility of heterogeneous interpretations of third person plural pronouns). Moreover, their position on the nominal segment of the Animacy Hierarchy is unclear, as third person pronouns may refer to human, animate and inanimate referents. It is true that, in discourse, third person pronouns are probably more often used to refer to humans (Dahl and Fraurud 1996: 56, via Kibrik 2011: 413), but – and may-

be partly exactly for this reason – only relatively rarely do languages use formally different pronouns to refer to humans and non-humans, or show any other third person split lower on the hierarchy. This is what English does by means of its *he ~ she ~ it* pronouns (the difference which is, however, lost in the plural) or Ket (Vajda, this volume, 2.2.1), but such splits seem uncommon. It is probably even much less common if one does not count cases like that of various East Caucasian languages, which make this distinction due to the combination of the fact that (a) demonstratives, when used headlessly for third person reference, show the same gender agreement as in their attributive uses, and that (b) their gender categorization is consistent with certain cut-off points on the Animacy Hierarchy, including a common distinction between human vs. non-human agreement in the plural. Notably, in East Caucasian, too, this correlation of third person reference with animacy is epiphenomenal, resulting from neutralization of the masculine and feminine agreement in the plural into human plural as opposed to non-human plural agreement. In other languages, the position of third person pronouns on the Animacy Hierarchy is variable and depends on the context (see also Dixon 1994: 90 on languages where proper names are assigned a position higher than demonstratives). Finally, in terms of their linguistic categorization, third person pronouns more often than not are attributive demonstratives different from personal pronouns proper in both morphological and morphosyntactic terms (Bhat 2003: 13–15, Bhat 2013, Siewierska 2004: 6), and may show divergent patterns of pluralization. Note that the same does not apply to the degree of individuation, which is one of the components of Croft's *extended animacy hierarchy* (2003: 128–132), because third person pronouns are similar to the first and second person pronouns in that they are necessarily individuated. We conclude that the position of third person pronouns with respect to personal pronouns and to nouns is ambiguous both conceptually and in terms of exponence of plurality.

To sum up, it seems safer not to assume any specific configuration of the upper segment to be crosslinguistically universal, as Fig. 9 effectively does. Therefore, for the sake of clarity, below we will only compare nominal plurals with personal pronouns of the first and second person.

Effectively, these considerations reduce the assumed "upper segment" of the Animacy Hierarchy to just one point on the scale: plural pronominal reference, with its internal hierarchical organization subject to cross-linguistic (first vs. second person) and contextual (third as compared to first and second person) variation. Bearing this in mind, let us now turn to the arguments for and against the unified view on (pro)nominal plural reference as represented in Fig. 9.

In section 3 above we have seen that there is an important difference between pronominal and nominal plurals in terms of heterogeneity vs. homogeneity. In most references to the Animacy Hierarchy, the fact that the nominal and pronominal plurality are conceptually different is effectively neglected. Corbett (2000) interprets the occurrence of heterogeneous plurality in terms of its being restricted to the left periphery of the Animacy Hierarchy, comparable to the effects of recategorization

of the nouns on the right periphery of the same hierarchy obtained in pluralization of mass nouns (such as sortal and abundance plurals). He also considers not only plurals of personal pronouns but also associative plurals (APL, 'X and his/her associates'), which provides a possible conceptual argument in favour of placing pronominal plurals in the same hierarchy as nominal ones. Heterogeneous plurality spanning personal pronouns, proper names, often kin terms and sometimes other nouns with human reference seems to be a case of a phenomenon whose scope includes both pronouns and the higher segment of the nominal part of the Animacy Hierarchy. APL is a plural of a noun but is heterogeneous in a way so similar to pronominal plurals that, as discussed in section 3, early references to APL equated it with pronominal plurality; see also Corbett (2000: 101–110, at least for the purposes of projecting both onto the Animacy Hierarchy).

However, Daniel (2020) argues that apart from the heterogeneous plurals there is scarce cross-linguistic evidence for positing proper names, and weak evidence for positing kin terms, as special cut-off points on the Animacy Hierarchy. This is in contrast with frequent morphosyntactic splits between human vs. non-human nouns, or animate vs. inanimate nouns, including plurality splits (Haspelmath 2013) but not limited to them. Daniel suggests an alternative explanation of the observed scope of heterogeneous plurals. APL are commonly derived from nouns whose referents are unique, most often from proper names and, less frequently, from kinship terms and highly individuated common nouns with human reference. This property of APL is shared with plural personal pronouns, because the speaker and the addressee arguably are also unique for the purposes of linguistic categorization – see Siewierska (2004: 82–83), Cysouw (2001: 70–73) for a convincing conceptual argumentation, and a typological study of second person plurals in Simon (2005). But it is only indirectly related to the dimensions of the Animacy Hierarchy such as animacy or degree of individuation.

Viewing unique reference as the key property of APL has the advantage of providing an explanation why pluralizing a proper name often leads to a reinterpretation as an APL ('N and associates'), even though the marker elsewhere signals additive plurality, which is a very common way to express APL crosslinguistically (Daniel and Moravcsik 2013). The use of the regular plural as APL marker may be explained as derived from the pragmatic incompatibility of regular additive plurality with inherently unique reference of a proper name (or high individuation of a topical NP in discourse, amounting to the same effect), thus resulting from its interpretation in context. For example, in Tswana, Southern Bantu (Creissels, this volume), the APL reading of a regular plural marker is the only one available for proper names; it is also possible with kin terms but not with other nouns.

The distinction between the expression of the APL by a dedicated marker as in Hup, Naduhup (Krasnoukhova, this volume) and West Circassian, Northwest Caucasian, on the one hand, and a regular plural marker, as in Ese Ejja, Pano-Tacanan (Krasnoukhova, this volume, 2.3.3), Tswana (Southern Bantu), Nivkh (isolate), Nga-

nasan (Uralic), Japanese (Japonic) and Eastern Dan (Mande), on the other, probably with intermediate cases of special syntactic or morphological constructions, as in Karko, Nilo-Saharan (Jakobi and Dimmendaal, this volume, examples (37)–(40), to become relevant later for the discussion of Fig. 11, is another telling example of the difference between grammatical and *interpretative* instantiation of a typological category discussed above for paucals in section 4.4. In languages of the type of Tswana, APL is not a linguistic category but an interpretation of a general plural marking.

Another case in point are *dyadic kinship plurals* in West Circassian (Northwest Caucasian). Dyadic kinship plurals are forms of kinship terms that render the meaning of 'people being in the named kin relation to each other', e.g. reference to a group consisting of a brother and a sister or of a mother and a child based on the nouns for 'brother' and 'mother', respectively (Evans 2006). This category is usually expressed by dedicated morphology (including in Nganasan, Uralic, and Gooniyandi, Bunuban, this volume). West Circassian, whose flexible morphology blurs the boundaries between lexical categories, uses the reciprocal suffix for this purpose, interpreted as a designation of reciprocal kinship relation between the elements of the group. This interpretative nature of expression also influences the scope of the category; unlike typical dyadic kinship terms attested in Nganasan and Gooniyandi, West Circassian forms can only be derived from symmetrical kin terms (thus, 'brothers (to each other)' from 'brother'), and also expand beyond kinship terms (thus, 'enemies (to each other)' rather than '(my several) enemies' from 'enemy').

Both associative plurals and dyadic kinship plurals are restricted to a specific segment of the Animacy Hierarchy as represented in Fig. 9. Both phenomena can be accounted for in terms of special semantics or pragmatics they are associated with, without involving the position on the Hierarchy as a functional explanation. Thus, dyadic kinship terms are a special type of plurality associated with kin terms; as mentioned above, kin group / family pronouns do exist (Corbett 2000: 26–30; Dalle Ceste, ms.) but seem to be much more rarely attested. As for APL, while uniqueness presumption and the Animacy Hierarchy are strongly correlated, Daniel (2020) provides some – though, we admit, marginal – evidence for linking heterogeneous plurals directly to the uniqueness presumption, such as APL-like plurals referring to sets of inanimate entities. Our interim conclusion is that the evidence from heterogeneous plurals, which, at first, seemed to provide a strong argument in favor of the unified approach, is not unequivocal.

Some of the properties of nominal plurals may provide evidence against the unified approach. In section 3 we have discussed that plurals of personal pronouns are capable of encoding 'pure' cardinalities ('2', sometimes '3', sometimes even '4', as in Russian Sign Language, Kimmelman et al., this volume), while cardinality is not in the focus in nominal number. Even when the nominal dual is present, it often expresses a special type of cardinality two, that of cohesive dyad. Attraction to cohesion in nominal duals cannot be explained in terms of the Animacy Hier-

archy as shown in Fig. 9, and leads to configurations like that in Tocharian (Indo-European), with a *cardinal* dual in reference to pure cardinalities in both nouns and pronouns and a separate *ambal* dual in reference to dyads for nouns only (see Kim 2018 for references); or an *ambal* dual for dyads in nouns but no dual at all for pronouns, as in Pilagá, Guaicuruan (Krasnoukhova, this volume, 2.3.3). This is in clear conflict with the unified view of the hierarchy, featuring more number values in nouns than in pronouns. Further, it seems that the approximative size values (paucal and greater plural) may be more common with nouns than personal pronouns. In Ket, Yeniseian (Vajda, this volume, 2.3.2), probably Marori, Trans-New Guinea (see discussion in Arka and Dalrymple, this volume, 2.6) and some languages of South America (Krasnoukhova, this volume) the paucal is only available to nouns; and in Kakataibo, Panoan (Zariquiey, this volume, 2.1), only to third person pronouns. All this shows that the interaction of personal pronouns and nouns with the category of number may be other than predicted by the Animacy Hierarchy.

We have seen in section 3 that the pronominal plurality, predominantly heterogeneous, is substantially different from the nominal plurality, predominantly homogeneous. Therefore, we cannot take for granted the relevance of the Animacy Hierarchy in Fig. 9 to the linguistic expression of number, and can defend an alternative "separatist" view on the pronominal and nominal number in terms of cross-linguistic disposition to express number distinctions, as shown in Fig. 10.
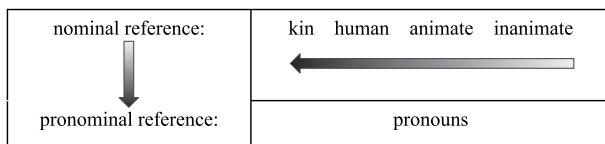
Fig. 10: Pro- vs. nominal plurality: a separatist view.

In Fig. 10, the cross-linguistic occurrence of number is shown along two independent dimensions: type of reference (vertical, between nouns and pronouns) and animacy (horizontal, for nouns) The decomposition of the Animacy Hierarchy along dimensions of *animacy proper* (human, animate, inanimate) and *referentiality* (pronoun vs. proper name vs. common noun) is proposed, on purely analytical grounds, by Croft (2003: 130), with a special status assigned to proper names based on their high topicality. Ivani (2017, 6), too, considers plurality splits across nouns and formal relatedness of nominal and pronominal plural marking in separate sections of her study (sections 6.3 and 6.5, respectively), reflecting the same intuition as conveyed by the representation in Fig. 10. In conceptual terms, Fig. 10 seems to be not less if not more compatible with what we know about the differences between the nominal and pronominal number than Fig. 9. The vertical dimension accounts for the preference of pronouns over nouns as a whole to show number distinctions. The

horizontal dimension accounts for the preference of nouns with referents more similar to humans to express number than nouns with referents less similar to humans.

With a view to the extensive discussion in Plank (1989), we would be cautious about making a similar hierarchy-based claim in relation to the *cohesive dual*, a category dedicated to natural dyads and expected to be associated with inanimate rather than with animate reference (except probably associative duals for married couples such as 'husband and wife'); more empirical research is needed here.

Let us now turn to formal empirical evidence that may support or refute the unified (Fig. 9) or separatist (Fig. 10) view on number. It has been noted that plural personal pronouns only rarely use the same formal means of pluralization as nouns (see Daniel 2013 for a crosslinguistic survey). In the languages of the volume, we observe shared (pro)nominal means for pluralization in Kakataibo (Zariquiey, this volume) and probably in Japonic (Shimoji, this volume; though see the discussion below). All other languages either show separate strategies of pluralization (as Indonesian, Malayo-Polynesian, see Gil, this volume) or at least stem suppletion together with regular pluralization. This kind of evidence is equivocal. On the one hand, one can see this as an argument against unifying pronominal and nominal number because most languages treat them separately, and shared marking may be accounted for by additional logic such as diachronic development. Indeed, in one relatively transparent case, the marginal use of the pronominal plural marker *-deih* on nouns (interpreted as APL) in Cantonese (Sinitic) is interpreted as erosion of the demonstrative stem, so that the construction Noun + Dem-Pl$_{pron}$, which is a common crosslinguistic type of periphrastic APL, is contracted to Noun-Pl$_{pron}$ (interpretation in Stephen and Yip 1994: 87). On the other hand, it may be viewed as an argument against the separatist interpretation, because, after all, some languages treat them together, and, under the separatist view, one has to explain how this is possible in the first place (probably, also relying on diachronic evolution of the systems). Modern Eastern Armenian (Indo-European) clearly separates pronominal and nominal plurals by using two different plural suffixes, *-k^h* and *-(n)er*. Historically, however, this results from the expansion of the innovative *-(n)er* (of unknown origin) into the nominal domain at the expense of the older *-k^h*, constraining the latter to the domain of pronominal plurality (de Lamberterie 1979). (Note that in both Cantonese and Eastern Armenian the markers in question are not dedicated to personal pronouns but are more generally used with various pronominal stems.) It is unclear how these paths of evolution should be interpreted in the light of the conflicting views on the two types of pluralities.

In some cases of shared plural marking, even the synchronic evidence may be subtle. In Standard Chinese (Sinitic), the pronominal plural suffix *-men*, obligatorily occurring on personal pronouns, is also marginally attested on nouns, almost exclusively nouns with human reference. No other productive pluralization devices are available to nouns. Moravcsik (2017) takes this as evidence in favor of the Animacy Hierarchy effects in Chinese: nouns more similar (closer) to pronouns on the Anima-

cy Hierarchy may borrow pronominal plural markers; and a similar interpretation is accepted in Daniel (2013) and in Ivani (2017: 15–16). However, based on corpus evidence, Iljic and Luo Yi (2001) suggest that the use of *-men* on nouns is associated with a specific use of NP, that of address. If this is true, NPs carrying *-men* have pronominal rather than nominal reference (refer to the addressee), which means that the suffix continues being associated with pronominal plurality even when used on nouns. Instead of being interpreted as occurring at the top of the nominal hierarchy shown in the first line of Fig. 10, these uses can probably be grouped together with pronouns in the second line.

The example of Standard Chinese makes it clear that we need not only to register cases of use of the pronominal plural markers on nouns but also to examine the functions the respective NPs have with such markers. Shared plural exponence in pronouns and nouns in Standard Chinese may be interpreted as resulting from including nouns into the scope of heterogeneous plurality, which would suggest an alternative interpretation of the position of the nominal associative plural forms in the typology of number independently of how these are expressed. In the separatist model, both the use of nouns in pronominal reference and associative plurals can be grouped with plural personal pronouns rather than nouns, as in Fig. 11, where arrows show the suggested dimensions of numerosity and dotted and solid lines showing most common patterns of shared exponence:
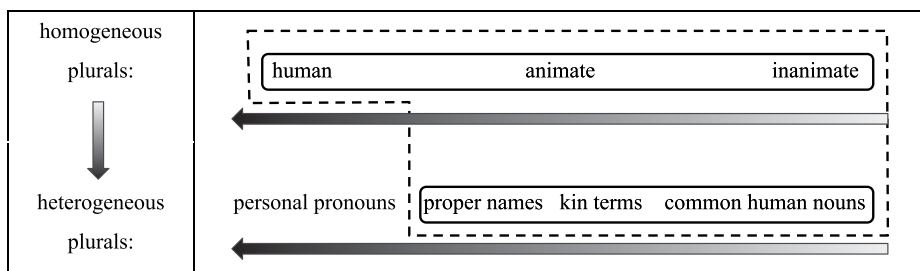
Fig. 11: Pro- vs. nominal plurality: an extended version of the separatist view.

In Fig. 11, the vertical dimension represents the distinction between heterogeneous and homogeneous plurals with a preference for the plural to combine with pronouns (necessarily heterogeneous plural) than to express homogeneous plurality more common for nouns. As suggested at the beginning of this section, this preference follows from the central role speech act participants play in communication, making quantification of pronominal reference very common crosslinguistically. It is in accordance with the common view on the Animacy Hierarchy.

What is different is that there are two separate horizontal dimensions for each type of plurality, one to account for heterogeneous plurals and the other for homogeneous plurals. Unlike Corbett's Animacy Hierarchy where these two types of nu-

merosity of entities are integrated with one another (Fig. 9 above), in Fig. 11 they are shown as parallel but independent hierarchies. To account for the occurrence of homogeneous plurals, we have the preference of nouns higher in animacy over nouns lower in animacy, as also shown in Fig. 10 and Fig. 9. What has been added in Fig. 11 is a preference for expressing *heterogeneous* plurals on pronouns over nouns and across different classes of nouns, reflecting the observation that most languages have plural personal pronouns, but quite a few languages do not have the APL category and in those that have APL its occurrence is by far more frequent with proper names and probably with kin terms than with common nouns with human and especially non-human reference. Occurrence of number marking along each of the hierarchies depends on the language. Standard English and quite a few languages have obligatory homogeneous number marking on various classes of nouns down the Animacy Hierarchy but have no grammaticalized expression of APL on nouns. Conversely, Maxakalí (Nuclear-Macro-Je) has singular and plural pronouns as well as APL in nouns but no homogeneous plurals are available for nouns (Nevins and Coelho da Silva 2020).

We can also think of similarities between the two dimensions. Whether a language has or does not have APL, the speaker can use a singular form of a proper name or a kin term in constructions like *X has arrived* implying that X arrived together with others but without making this explicit in the linguistic form. This heterogeneous interpretation of a form unmarked for number can be compared to the use of number-neutral forms of nouns to refer to more than one entity in languages where grammatical expression of the plural is not obligatory or is unavailable (see section 4.2). We suggest that the availability of number marking along the two scales is in principle independent, which means that the two scales cannot be combined with one another.

How does Fig. 11 account for formal expression of numerosity? Here, the relevant distinction is that between the two most common types of expressing associative plurals, one that uses the same marking for APL (heterogeneous) and additive (homogeneous) plurals, shown as dotted line, and the other one that formally distinguishes between additive plurals and APL, shown as two solid lines (Daniel and Moravcsik 2013). Typically, pronominal plurals are expressed separately from both, e.g. by number suppletion or dedicated morphology. Apparently, the dotted line does not comply with the suggested model because it spans both scales, that of homogeneous and that of heterogeneous plurals. Tentatively, we suggest that the explanation lies in the general tendency of pronouns towards special plural exponence, as supported by their separation also in the systems with dedicated APL. In other words, we interpret the dotted line as showing a generalized plural covering both homogeneous and heterogeneous plurals with the exception of irregular expression of heterogeneous marking on pronouns. Nevertheless, we admit that, before a thorough survey of all formal patterns is carried out, this may be seen as a strong argument in favor of the mono-dimensional view of the Animacy Hierarchy.

In Fig. 11, the case such as the use of Standard Chinese *-men* on nouns for address, is subsumed under personal pronouns. This exact pattern is so far uniquely attested in Chinese. However, even apart from being a frequent lexical source for pronouns, the use of lexical nouns in address or self-reference may not be as cross-linguistically marginal as it may seem (Siewierska 2004: 2), so one may look forward to discovering more such evidence. In Archi, Dargwa and Ingush (all East Caucasian), nouns used in such contexts switch to the gender-number agreement pattern unique to plural personal pronouns. Further, in Japanese the use of the pronominal plural marker *-tati* is obligatory only with what Shimoji (this volume) calls Address Terms, a class of nouns used in both (a) addressive (second person) and (b) referring (third person) functions. With these nouns but not other nouns, it has associative plural interpretation. While (a) makes it similar to the Chinese *-men*, (b) looks more like a true expansion of pronominal plurals into the nominal domain of pluralization. As a potential evolutionary path of the Japanese pattern, one can suggest that the use of the marker could have started from the addressive function only – similar to, and possibly supported by the influence from, Chinese – and then shifted to other uses of the same nouns (cf. the discussion of *vocativus pro nominativo* development in Stifter 2013), but preserving their preference for heterogeneous reading.

Tentatively, one could also suggest the possibility of an opposite scenario for synchronically shared patterns of pluralization: in the languages with heavy use of nouns for personal reference (Siewierska 2004: 247–249), nominal plural marking may spread to the domain of pronominal plurality not by broadening their semantic range to include heterogeneous reading but by nouns developing into personal pronouns, bringing along their plural marking into the domain of pronominal plurality. In both hypothetical directions of expansion of plural marking, formal extension comes in the wake of conceptual extension, opening a way to model the expansion of pronominal plural marking to nouns or vice versa without a need to posit a single scale where pronominal plurals are adjacent to human plurals. We conclude that the attestation of shared marking cannot be used as an argument either against or in support of the separatist approach. What it does is it highlights the importance of diachronic study of the evolution of (pro)nominal number marking.

To sum up: we do not argue against the vast cross-linguistic evidence confirming that animacy may account for the scope of number marking among nouns, nor do we ignore the fact that personal pronouns display a much higher propensity for being marked for number. However, we cautiously cast some doubt on the practice of integrating these two facts into a unified Animacy Hierarchy. The two preferences can alternatively be explained by two separate functional (or cognitive) types of pressure from human communication: more relevance attributed to the numerosity of speech act participants than to the numerosity of other referents (hence pronominal plurals being more crosslinguistically common than nominal plurals, and restricted spread of heterogeneous plurals possibly involving some nouns) and more relevance attributed to the numerosity of humans / animates than to non-humans /

inanimates (hence relevance of animacy for nominal plurals, and restricted spread of additive plurals down the Animacy Hierarchy). Too few empirical facts require that personal pronouns be placed on the same scale as nouns, and some indicate the opposite direction (more focus on cardinal numerosity in personal pronouns; possibly, greater availability of approximative numerosity in nouns). We leave this open until more evidence in favor or against the model that places pronouns and nouns in one dimension to build a single scale becomes empirically available.

6 Verbal number

The expression of numerosity associated with the verb, also known as *verbal number*, has become an important topic of typological research starting from Durie (1986); about the same time, and independently from him, Dolinina (1989, 1990) made a number of insightful observations on (a)symmetries between nominal and verbal number. Of the now relatively rich literature, we should also mention early treatments of event numerosity in Dressler (1968), Cusic (1981), a most recent and detailed typological survey in Mattiola (2019, 2020); and in between, influential discussions in Corbett (2000) and Cabredo Hofherr and Laca (2012).

Following Newman (2012), Mattiola (2020) departs from the use of *verbal number* in Corbett (2000) and opts for *pluractionality* in the sense that is inclusive of Corbett's *participant number*. However, as we argue below, the term *pluractionality* does not naturally apply to the contexts which do not involve numerosity of events, i.e. to those uses of Corbett's *participant number* which are not distributive. In this volume, we stay with the terminological choice made by Durie and Corbett, i.e. *verbal number* as a general designation of verbal expression of numerosity.

Across the languages in the volume, verbal number is attested in Karko (Nilo-Saharan) and Eastern Dan (Mande) in Africa, Arabic (Semitic) in Near East, Nivkh (isolate) and Ket (Yeniseian) in Northern Asia, Indonesian (Malayo-Polynesian), Murik (Lower Sepik) and Marori (Trans-New Guinea) in the Pacific, Gooniyandi (Bunuban) in Australia, Kiowa (Kiowa-Tanoan) and Mohawk (Iroquoian) in North America, in various languages of South America (including Kakataibo, Panoan represented by a dedicated chapter), and in Russian Sign Language. This is a rich representation that cannot be attributed to differences in the interpretation of the category by different authors, most of whom explicitly discuss verbal number as distinct from verbal aspect. Thus, the list above does not include the languages where the domain of event numerosity is only represented by e.g. the iterative category. It is also noteworthy that at least in Marori, Ket, Karko, Eastern Dan, Kiowa and Kakataibo the category is often instantiated by stem suppletion, a known signature property of verbal number. In short, whether by chance or reflecting the actual cross-linguistic spread of this category, verbal number is well represented.

In areal terms, Krasnoukhova (this volume) makes a substantial amendment to the cross-linguistic coverage of verbal number by providing a detailed account of verbal number in South America. It is true that the languages of South America have been mentioned in typological discussions of verbal number, including 29 languages in Mattiola 2019, but only Kaingang (Nuclear-Macro-Je) in Durie 1986 and only Xokleng (Nuclear-Macro-Je), a close relative of Kaingang, in Corbett (2000: 249). The area deserves a much more detailed study, as it turns out to be a most important hotbed of verbal number, found in 35 out of 41 genetic groupings in Krasnoukhova's survey (this volume, see also for a general survey of the sources). As one additional highlight, quite a few languages of South America feature more than a two-way opposition in verbal number, a pattern which is thought to be rare elsewhere (Corbett 2000: 250).

Note that a related topic that was not addressed in the questionnaire and, consequently, is not necessarily covered by the contributions to the volume, is the domain of lexical verbal quantification, i.e. the expression of meanings such as 'twice', 'regularly' and 'rarely' outside the verb itself (for the discussion, see e.g. Birjulin 1989, Gil 1993), even if some contributions may briefly mention it (e.g. McGregor, this volume, 2.4; Fassi Fehri, this volume, 4.3).

The main impact of Durie (1986), and the first step towards understanding verbal number as a category on its own, was problematizing the notion of verbal agreement by suggesting that some ways of conveying the information on the number of participants on the verb are better interpreted as changing the lexical meaning of the verb so that it *selects for* rather than *agrees with* a plural instead of a singular participant (see also Mithun 1988, 1999: 84–86, Mattiola 2020) than as agreement controlled by an argument. The claim was based primarily – but not exclusively – on the cross-linguistic attestation of formally unrelated verbal stems for different number values of participants traditionally viewed as stem suppletion for number (see Veselinova 2006, Chapter 7 for a typological survey). Depending on the language, there may be clear arguments in favor of distinguishing verbal number from agreement. First, the choice of the alternating stems in most languages reflects the number of S/P arguments, resulting in what appears to be an ergative pattern, including those languages where agreement is controlled by S/A arguments, instantiating an accusative pattern (see early discussion in Frajzyngier 1985, Durie 1986, and Mithun 1999: 84, and Krasnoukhova, this volume). Second, formal differences associated with the number of participants and reflected in the use of different stems may only involve few verbal meanings, while other verbs do not show agreement at all. An overview of verbal meanings typically involved, including verbs of motion and (change of) posture verbs, is provided in Corbett (2000: 257–258) and Veselinova (2006); see also Krasnoukhova (this volume, 2.4). Whether one accepts the specific point about verbal number affecting the lexical rather than grammatical meaning, and the interpretation of its morphological exponence as derivation rather than inflection, Durie's arguments lead to the inescapable conclusion that express-

ing the number of participants may be an inherent category of the verb rather than merely a morphosyntactic copy of nominal number (see also Frajzyngier 1985). What Corbett (2000) further contributes to the theoretical discussion is arguing for distinguishing two separate types of verbal number, *participant number* vs. *event number*, the latter category expressing the numerosity of events.

The distinctions between number agreement, verbal number of the participant number type and verbal number of the event number type may seem conceptually clear, but they get blurred when examining individual systems, as pointed out in Dolinina (1989), Mattiola (2020). Corbett notes that the plurality of a participant often implies plurality of events in English (Germanic) which lacks the category of verbal number, and the plurality of events often implies plurality of participants in Pomo (Pomoan) without number marking on NPs being involved (Corbett 2000: 244). He also discusses diachronic border cases of Gitksan (Tsimshian) and Huala-pai (Cochimi-Yuman) where the verbal number has almost reached the point where it can be interpreted as number agreement (Corbett 2000: 256); and so does New-man (2012, 4.4). Diagnostics for delimiting different phenomena are discussed in detail by Durie (1986) and reviewed or recapitulated in Corbett (2000: 252–255) and Mattiola (2020), but some have been suggested already in Frajzyngier (1985). The contributions to this volume strive to be clear as to what type of verbal number they deal with, but sometimes the evidence is simply not sufficient. It is not very clear to what extent this comes from deficient evidence or from the fluidity of the category itself – probably both. Thus, Corbett (2000: 249) suggests the label 'mixed' for the verbal number ambiguous between event and participant number (cf. Fassi Fehri, this volume, 2.4.2). Krasnoukhova, in one of the most explicit discussions of the distinction in this volume, suggests that the languages of South America vary along the dimension of event vs. participant number not in a categorical way, as partici-pant vs. event vs. mixed type, but rather in terms of gravitating towards one or the other type. In addition to this fuzzy distinction between participant and event num-ber, each of the two categories has to be distinguished from their own functional neighbours on both sides, the former from agreement and the latter from the aspec-tual domain (iteratives and multiplicatives), as shown in Fig. 12.

Morphosyntax	Verbal Number		Aspect
number agreement	participant type	event type	multiplicative, iterative etc.

Fig. 12: Verbal number and its functional neighbours.

Distinguishing between number agreement and participant number is based on evi-dence that the latter is a morphological category whose occurrence is partly or even fully independent from the nominal number. Radical discrepancies happen in lan-guages where the number of participants can only be expressed on the verb, and

there is no nominal number at all. Elsewhere, verbal vs. nominal number may have different sets of values. In some cases, one could argue that verbal agreement under-differentiates number values, as in Bayso (Cushitic), where both nouns marked for paucal and nouns marked for plural control plural agreement (Corbett 2000: 10–11; but see Corbett 2006: 172–174 for more details). But if verbal but not nominal number has a dual, (Durie 1986, Corbett 2000: 255, Krasnoukhova, this volume, 2.5), or if verbal but no nominal number is attested as a category in the first place (Nevins and Coelho da Silva 2020), one is forced to admit verbal number as an independent category.

There is a range of cases of milder mismatches between the two categories. For Gooniyandi (McGregor, this volume), the expression of number in the NP and on the verb is controlled by different factors (such as group cohesion), which, however, often produce similar outcomes and may thus be mistaken for agreement. Bagirokova et al., this volume, focus on predicate – argument number mismatches in West Circassian, where they are, however, less principled than in Gooniyandi, and remind us of the yet more marginal cases of mismatches observed in English where a handful of nouns, like *fish*, do not express number morphologically but may agree in the plural, and the singular of some group nouns, such as *committee*, may combine with plural agreement on the verb. It is not clear how radically the cases like Kiowa are different from Gooniyandi, Gooniyandi from West Circassian, and West Circassian from English; in the latter case at least, the difference seems in the degree of mismatches and probably depends on how much emphasis is put on them in the descriptive discussion.

expression of number on the predicate	absent	present			
			Verbal Number		
	(a) number expressed on nouns but not on verbs	(b) values of number on the verb copy the values of number on its arguments (*canonical agreement*)	(c) same values of number on nouns and verbs, contextual mismatches (*semantic agreement*)	(d) different inventories of values of number on nouns and verbs	(e) number expressed on verbs but not on nouns
		Nominal Number			
expression of number on the argument		present			absent

Fig. 13: Expression of the number of participants on the predicate and on the argument.

Fig. 13 shows various possible configurations of different loci for the expression of the number of participants. Thus, if it is expressed on the argument itself and cannot be expressed on the predicate (a), including by ways of agreement, it is neces-

sarily interpreted as nominal number; but if it is expressed on the predicate and cannot be expressed on the argument (e), as in Karitiâna, Tupian (Krasnoukhova, this volume), it is necessarily interpreted as verbal number. There are also configurations, shown here as (b) to (d), where the number of participants is, or at least can be, expressed on both the noun and the verb and one must decide whether one is dealing with two independent categories. A relatively clear case is that of (d), with the number of participants displaying different inventories of values on the predicate and on the argument. The case in (c) is less straightforward: the question whether we are dealing with two imperfectly aligned independent categories, nominal and verbal number, or with number agreement on the verb which is an imperfect morphosyntactic copy of the nominal number of its controller becomes the matter of theoretical stance and researcher's focus. Note there is a certain asymmetry in Fig. 13: whenever participant number is expressed on the noun, it counts as an instance of nominal number, while some cases where participant number is expressed on the predicate – more specifically, (b) – do not count as an instance of verbal number. In principle, the situation shown in (b), with the values expressed on the predicate strictly identical to the values expressed on the argument, could be considered not as morphosyntactic agreement but as expression of participant numerosity by means of verbal number that happens to be perfectly harmonic with the expression of participant numerosity by means of nominal number. This is, however, not the case in theoretical discussions and descriptive grammars: our conception of the number of participants departs from the assumption that its default locus of expression is the NP. In part, this may reflect our conventional ways of thinking of grammar but there is more to it than just a convention (see below on asymmetries between participant and event numerosity). Note that based on the same assumption, one could also account for (d) in terms of agreement by positing patterns of syncretism of the expression of number values on the verb or on the noun, but again, this is not what is usually done. In any case, the situation in (c) remains intermediate and may obviously shift leftwards towards agreement or rightwards towards verbal number depending on how prominent are the observed mismatches in the language grammar, and how much emphasis is put on them in the study. See further discussion in Mattiola (2019, 2020).

The fact that the expression of the participant number on the verb may be independent of the nominal number leaves one with a question: is this merely a matter of the locus of expression of the same concept or a related but distinct category?

In some cases, number mismatches are a grammaticalized means for conveying a specific plural interpretation. In this volume, McGregor and Bagirokova et al. discuss in depth several such cases, which can be very subtle. Corbett observes that plural agreement with a singular NP may convey associative plural in Maltese (Semitic) and in one of the northern dialects of Russian, Slavic (2000: 191–192), and singular agreement with a plural NP is a means to convey the dual in Pomo, Pomoan, and Hopi, Uto-Aztecan (Corbett 2000: 169), the phenomenon that Corbett

introduces as *constructed* number values (with an earlier mention in Plank 1989). On the basis of the data in the volume we can only tentatively suggest that, in cases of mismatches, verbal plurality in combination with singular NP tends to express a distributive reading and nominal plurality in combination with singular verb a collective reading. That verbal number profiles distributivity as opposed to nominal number, which is more often than not neutral in this respect, is in accordance with Mithun (1988) who argues that the grammaticalization of nominal number that takes verbal number as its point of departure is often based initially on distributive numerosity. It has also been noted that distributivity over participants may lean towards exhaustivity, focusing on the fact that every single element of the set represented by the argument is affected by the situation (Dolinina 1989, Vydrin, this volume, 2.4.1).

As a side note, number mismatches between the argument NP and the predicate seem to be mostly relevant in the case of agreement with full argument NPs and not with independent pronouns of any person. In pronouns, the expression of number in the first and second person is overwhelmingly obligatory; in the third person, if available, it is also most often obligatory. One case of morphologically singular pronouns with morphologically plural verbs in this volume is discussed by Bagirokova et al. (this volume, examples 106–108) and is interpreted as general number in third person pronouns.

Once we have ascertained that the relevant means of number expression on the verb is not agreement, it is further necessary to distinguish it from the *event numerosity*. Indeed, multiplicity of participants can easily imply multiplicity of events, and the other way around. If a predicate describes a situation involving a non-singular participant – a group of people or a collection of objects – it can be viewed as a single collective event involving all elements of the set together or as several events with each element of the set separately involved. Then, a possible test for the type of verbal number one is dealing with is whether a verbal predicate marked as plural is compatible with a single participant at all (cf. Mattiola 2020, and implicitly in previous discussions). If it is not, we may conclude this is an instantiation of participant number. If it is, we may assume this is an instantiation of event number, which in this case (i.e. when combined with a singular participant), necessarily leads to an interpretation distributed over time and possibly also over space. It is difficult to imagine a situation with a single participant where the event is distributed over space without being distributed over time. One can only think of something simultaneously happening to an entity at different places of its extension, such as a house burning at different locations, so a necessary prerequisite is that this participant is viewed as having a spatial extension. It is noteworthy that Birjulin (1989) includes distribution over time as a necessary component of distribution over participants (this prominence of distribution over time in his discussion of distributivity may be motivated by the context of the volume which as whole is focused on iterativity; see Xrakovskij 1987, 1989); and Dolinina (1989) only mentions distributivity over

participants vs. distributivity over time but not over space as subtypes of event numerosity. While distributivity over space remains an important interpretation of event number (Dolinina 1990, Mithun 1999, 3.2.5), we conclude that if the situation is not distributed over a set of participants, it is almost inevitably distributed over time.

In other words, the signature property of the event number is that it requires distributivity over some parameter of the situation, be it a participant or time or space (with distribution over space strongly implying distribution over time). We should expect that such a category would be compatible with distribution over any of the participants of a predicate with several arguments; and that it is incompatible with a collective reading of a non-singular participant, at least when the event that involves a collective participant is not distributed over time or space. These are expectations to be checked empirically, as the currently available data are insufficient.

Mutual implications connecting distributivity over participants, over time, and over space bring us back to the original Fillmorean (1968) view of semantic roles, in which temporal and spatial parameters of the situation were considered on a par with core arguments. A possible outlook for empirical research is the issue of the specification of distributivity selecting specific types of argument, such as P, or S, or non-core arguments of the verb, or various combinations thereof, or, finally, instantiations of the distributive categories encompassing not only participants in the narrow sense but also time and space. In Karko (Nilo-Saharan), stem suppletion is sensitive to the numerosity of S(ubject), P(atient), T(heme) and E(xperiencer) (Jakobi and Dimmendaal, this volume). In Kakataibo, Panoan (Zariquiey, this volume), the conceptual space is formally highly differentiated, with distributive, iterative, plural object and several other number-related meanings all expressed by distinct suffixes.

A possible classification of verbal number is shown in Fig. 14.

Participant number: Numerosity of Participants	Event number: Numerosity of Events
collective	distributed
	over participants over time
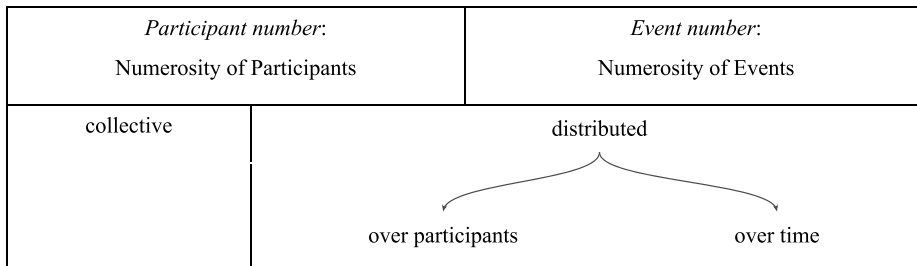
Fig. 14: A possible classification of verbal number.

Fig. 14 suggests that, based only on the contexts where a situation is distributed over multiple participants, we cannot decide whether we are dealing with partici-

pant number or event number; it is by considering other available uses that we can decide. In accordance with the view presented in Fig. 14, Corbett's 'mixed' verbal number must alternatively allow for collective and distributed plural participants indifferently rather than only for both plural distributed participant and distributed eventuality, because contexts with distributed participants may be interpreted as involving distributed eventuality. An open empirical question is how often participant number marking is dedicated to non-distributed, collective participants (cf. Karko, Jakobi and Dimmendaal, this volume), but it seems clear that participant number at least does not have to be restricted to distributed participants and is often indifferent with respect to distributivity vs. collectivity, just as nominal number most often is. It is clear, then, that contrary to Mattiola's (2020) suggestion, *pluractionality* cannot account for the whole domain of nominal number: participant number needs not be distributed and thus it does not necessarily entail numerosity of events, which is hardwired in this term.

As a final note on the delimitation of participant and event subtypes of verbal number, we point out once again that Krasnoukhova (this volume) treats the distinction as gradual, suggesting that, in a specific language, the category of verbal number that qualifies as 'mixed' number may gravitate towards participant or event number interpretation; and so does Corbett in the first place (2000: 249–250). Vydrin (this volume) comments that, in Eastern Dan (Mande), participant numerosity and event numerosity may combine in one context. Cross-linguistic availability of such ambiguous categories is not accounted for by the categorical representation of the distinction in Fig. 14, so one has to keep in mind that it is a simplification.

Dolinina (1989) notes that availability of readings involving multiple occurrences of an event along the time axis (i.e. distributivity over time) is what distinguishes event numerosity from participant numerosity. The same participant repeatedly participating in a series of situations remains the same participant, while the situations in which he or she participates may be viewed as different occurrences simply because they are distributed over time. Event number distributed over time borders on the domain of aspect, specifically on the domains of the categories of multiplicative and iterative (and still farther away, the habitual), which again raises the issue of delimitation.

The iterative and habitual categories are different from the event number, which may also be distributed over time and thus imply time-sequencing, in a way that is similar to how the imperfective aspect differs from the perfective aspect. The iterative/habitual describes the same situation that happens repeatedly over unbounded time (or has an unspecified number of iterations within indicated temporal boundaries, as in *Last year he was coming here almost every day*), with the iterative focusing on the repetition of the event (Xrakovskij 1987, Birjulin 1989) and the habitual representing the situation of the repeated event as a property of its participant (Dolinina 1989). Often, such descriptions provide a background to the main storyline. Event number, on the other hand, describes a bounded series of events with a definite but

not necessarily specified number of occurrences. To put it informally, an iterative conveys the idea of the situation repeating itself over (a period of) time, which implies but does not profile the idea of a certain number of repetitions. Event number, on the other hand, conveys the idea of the situation that occurs a number of times, which often implies sequence in time but does not profile this idea and may be distributed over other variables. When distributed over time, event number indicates that the situation occurs repeatedly within one time frame, and constitutes a single event with respect to the narrative viewed as a flow of events (see discussion of *verbal plurality* in Cusic 1981).

The last distinction we are going to make is between event number and the category of multiplicative which is also built on a numerosity of repeated subevents, such as designating situations of coughing, breathing, beating, chopping etc. by derivation from designations of a single such event. Similarly to event number, a multiplicative presents a sequence of situations as a single event in a flow of other events, but they may be different in the degree of integration of these situations into a single event. Informally, multiplicatives build a single event from elementary micro-events, and event number builds a single macro-event from elementary events, but this distinction is most probably fuzzy, depending on the verbal semantics. Therefore, the only clear distinction between event number on the one hand and both the multiplicative and the iterative on the other is that event number can – though does not have to – be distributed over variables other than time, is essentially distributive and does not necessarily imply sequence in time, while iteratives and multiplicatives necessarily refer to a sequence in time and, in this sense, belong to the temporal aspectual domain. This is schematically shown in Fig. 15.

Verbal number:	*Aspect*:
event number	iterative, multiplicative
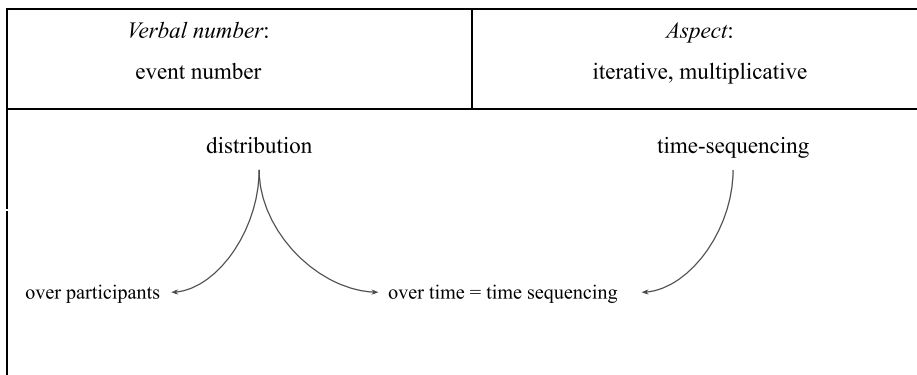
Fig. 15: Verbal number and its aspectual cousins.

Note that while different approaches may use the terms *multiplicative*, *iterative* and *habituals* in different ways, their interpretations seem to be aligned along the scale of temporal boundedness from the most bounded multiplicatives to the least bound-

ed habituals. For an attempt to integrate event number and these aspectual categories into one theoretical framework see Birjulin (1989). While he focuses on distributivity over time and on comparing lexical event quantification (such as 'twice', what he calls *counting* of events) and iteratives proper, he makes an interesting observation that allows a parallel between nominal and verbal number. In his discussion of multiplicatives and semelfactives, he notes that different directions of derivation may reflect whether it is the complex event or its elementary sub-event that is conceptualized as more basic. This is reminiscent of the tripartite and inverse systems of nominal number marking as discussed in section 2.2 above (see also Jakobi and Dimmendaal, this volume, Harbour and Mackenzie, this volume).

It is indicative that the studies that approach verbal number from the aspectual domain often define event numerosity (Šluinskij 2006) and verbal distributivity (Birjulin 1989) primarily or even exclusively in terms of distributivity over time (time sequencing), and do not discuss other possible types of event distributivity. Xrakovskij (1989) does not distinguish distributed and non-distributed plural participants, subsuming both under *distributivity*. From the opposite perspective of participant (vs. event) number, Mattiola (2020) groups most of the aspectual uses into 'additional functions' of verbal number, and keeps only iterative and frequentative as its core meanings. Also see Mattiola (2020, section 6) for a discussion of the impact of approaching verbal number from different perspectives.

We thus maintain that the distinctive property of event number is its distributivity (Mithun 1988, 1999: 88–91, further supported by some evidence in Mithun, this volume, Gruzdeva, this volume, among others), be it distributivity over participants, making it similar to participant number, or distributivity over time, making it similar to time sequencing in multiplicatives and iteratives, or distributivity over space. Distributivity over space may be implied by distributivity over participants because different participants cannot be at the same place simultaneously. Distributivity over space may itself imply distributivity over participants because the same single participant cannot be at different places simultaneously (cf. Mithun 1999: 89) or distributivity over time because the same single participant cannot be at different places at the same time. It can thus itself be the point of departure for other types of distributivity. Given that the varieties of distributivity (over participants, over time and over space) are so intertwined in language use, it comes as no surprise that, while discussing various types of distributivity makes sense for the purposes of cross-linguistic comparison (etic perspective or comparative concepts), individual languages may associate the category of verbal number with distributivity as a unified concept, with different 'subtypes' of distributivity being nothing more than its realizations in contexts. Moreover, dedicated means expressing distributivity over time alone may not easily be distinguished from aspectual categories.

Let us now come back to the comparison of nominal number and verbal number in semantic terms. The primary distinction in the domain of nominal plurality is the distinction between singular entities and sets with a cardinality higher than one, a

highly abstract notion. In some contributions that discuss verbal number, however, it transpires that verbal number may convey a less abstract notion of event *multiplicity*, the meaning of being many, i.e. a number of events that somehow exceeds pragmatic expectations (cf. the example of *murder* vs. *massacre* in Durie 1986, Mithun 1999: 85, Corbett 2000: 259; and the discussion of reduplication in Vydrin, this volume), as opposed to the meaning of being a few, or not many (cf. the discussion of non-plural verbs meaning 'one or two' events or participants in Foley, this volume, 2.3). In such cases, one may ask whether being many is a semantic extension of being more than one or is in fact the core meaning of verbal number, as discussed by Corbett (2000: 248, 250–251). This is definitely a possibility in systems described as having verbal number, although the lack of clear evidence bearing on the issue does not allow for a stronger conclusion. In the present volume, it is also noted that verbal duals may be interpreted as verbal paucity (or 'several-ity') rather than specifically the event cardinality of two.

As these meanings are attitudinal, one may further question how obligatory the category of verbal number is in these languages; and even outside this pragmatic nuance of event numerosity, it is unclear how obligatory this category is in each specific case. How often is the distributivity of events obligatorily reflected on the verb? The presence of marking of event numerosity entails distributed reading but does absence of such marking entail non-distributed reading, and if it does, to what extent? One cannot exclude that event and participant number might be different in this respect, with a tendency for participant number to be more obligatory than event number (see Corbett's 2000: 246 note on Hausa, Chadic branch of Afro-Asiatic); Corbett even takes obligatorification of participant number as one indication of verbal number evolving into number agreement (2000: 256–257). As Dolinina (1989) notes, numerosity of events is a much more fluid concept than numerosity of entities. Viewing multiple events distributed over time (and, we should add, over participants or space as well) as many events or one event depends on the perspective the speaker chooses much more strongly than viewing a set of entities as many entities or one single collective entity.

In short, we hope that future language-specific studies will pay special attention to how readily a verbal plural may refer to two events, and how readily a verbal singular can refer to more than one event.

Turning now to the formal expression of verbal number, out of the three formal means of expressing this category, namely, stem alternation, reduplication and affixation, at least the first two seem to be more closely associated with one of its semantic subtypes. Based on observations in Durie (1986) and supported by the survey by Krasnoukhova (this volume), there is a bias in the use of reduplication towards event number and in the use of stem alternation towards participant number. On Krasnoukhova's map, affixal expressions of verbal number are equally prone to convey either meaning; see Fig. 16.

While these form-meaning mappings need be confirmed on a larger sample, a tentative functional explanation for these asymmetries is that, following Durie, an

stem alternation	affixation	reduplication
participant number		event number

Fig. 16: Form-meaning associations in the domain of verbal number.

event involving many participants may be construed as a special collective event rather than a sum of sub-events, and thus leans towards a separate lexical encoding (i.e. towards using a different stem). On the other hand, reduplication may be iconically associated with distributivity, a conceptual connection also noted for reduplication as a means of expression of nominal number (e.g. Krasnoukhova, this volume, Kimmelman et al., this volume, Gruzdeva, this volume).

An important empirical contribution of this volume is that both trademark means of expression of verbal number, reduplication and stem suppletion, are attested with property words. This is not surprising because property words are predicates and verbal number strategies could thus be expected to extend to adjectives. It is possible that the expression of number on adjectives may be mistaken for agreement, and more attention should be paid to the semantic nuancing of this category. In fact, distributivity of the type 'various and assorted' in Mithun (1999: 89, 99) and Mithun (this volume), such as diversity in property realization distributed over the referents of which this property is predicated, may be a separate and less discussed kind of distributivity, distributivity over types. For a brief mention of the possibility of heterogeneous verbal numerosity and its similarity to the sort plurals in nouns, see Dolinina (1989).

A Tupían language Gavião and an Arawakan language Tariana (Krasnoukhova, this volume, 3.2), Marori, Trans-New Guinea (Arka and Dalrymple, this volume, Note 8), Kiowa, Kiowa-Tanoan (Harbour and McKenzie, this volume, Table 4), Eastern Dan, Mande (Vydrin, this volume, 3.1.2) have suppletive plural forms of some but few adjectives, eventually restricted to only one property word, 'small' in Karko, Nilo-Saharan (Jakobi and Dimmendaal, this volume, 2.3.7; incidentally, the same property word is also number-suppletive in e.g. Lak, East Caucasian), resembling alternating stems in verbal number.

As a probably more compelling example, Karko (Jakobi and Dimmendaal, this volume) features adjectives that agree in number (as well as non-agreeing adjectives). In formal terms, the expression of number on agreeing adjectives is relatively independent from its expression on nouns, being predominantly by suffixation on nouns and predominantly by tone alternation on adjectives. But it is also different in a more substantial way. In Karko, the plural of nouns may be formally marked or unmarked as compared to their singular forms, and there are also transnumeral nouns that do not inflect for number at all. Irrespective of the marking of number on the noun, adjectives take plural forms based on the reference of the head noun, not on its form (marked for plural or unmarked plural or transnumeral). This sug-

gests that adjectival number in Karko is not agreement but a separate category. Tentatively, these two instances of the category of *adjectival* number, distributivity over types discussed by Mithun (1999) and agreeing adjectives in Karko, correspond to event number and participant number on the verb, respectively, if property ascription is taken as analogue of eventuality. One must however note that, in Karko, not only adjectives but also some other attributes such as demonstratives behave in a similar way.

In several cases, the expression of number on the verb has the same formal exponence as the expression of number on nouns (West Circassian, Northwest Caucasian and Nivkh, isolate; see also Krasnoukhova, this volume, Note 20). It is unclear whether this correlates with any typological parameters of number expression on the verb, and both West Circassian and Nivkh seem to be closer to conventional agreement than several cases of verbal number described in the volume. On the other hand, this polyfunctionality must be indicative of the diachronic origins of number marking on the verb, with the marking of verbal number probably developing from the plural marking in the NP, in particular from the use of action nominals as nominal predicates or from participles acquiring finite uses (see Malchukov 2013, 4.5 on the origins of finite forms in Nivkh). This direction of evolution is opposite to the one described in Mithun as the extension of the formal scope of distributive marking from verbs to nouns and its further evolution into a means expressing non-distributive numerosity of entities (1988; 1999: 90).

| Nivkh (isolate), West Circassian (Northwest Caucasian) (this volume) | Number on nominals | predicative use of action nominals / participles $\longrightarrow$ | Number on verbs |
| Kwakwala (Wakashan), Quileute (Chimakuan) Mithun (1988, 1999) | | $\longleftarrow$ distributive number on nouns   verbal number | |
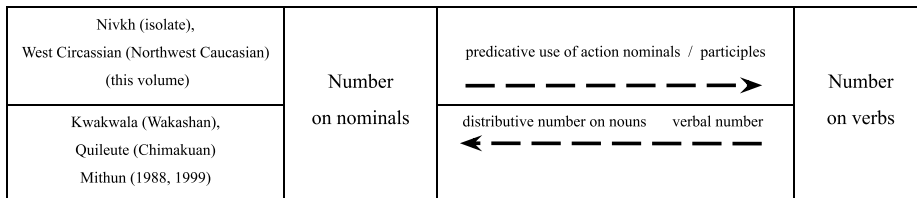
Fig. 17: Opposite directions of the evolution of number marking.

More generally, how related are the expressions of event and participant numerosity crosslinguistically? Among several types of asymmetries between nominal and verbal number, Corbett (2000, section 8.8) observes that verbs can express participant numerosity, but nouns cannot express event numerosity. He tentatively explains this asymmetry by the available paths of grammaticalization: verbal number may develop from the expression of nominal number, but itself cannot develop into nominal number (2000: 262–263). The latter is, however, the evolutionary path suggested by Mithun (1988) for some languages of North America as expansion of distributivity from events to participants (see Fig. 17). At the same time, Corbett's asymmetry is probably supported by the fact that while nominal morphology relatively often has a dedicated means for expressing collective plurality, a dedicated means

to express distributive plurality on nouns is relatively rare, and those that are attested (as distribution over types, meaning 'various and assorted') do not imply numerosity of events. If not from the universals of grammaticalization, the asymmetry may follow simply from the fact that participants make an integral part of the situation, and numerosity of participants affects the conceptualization of the situation (Mithun 1999, 3.2.4). On the contrary, the participant remains the same independently of the number of events he is involved in. Slightly rephrasing Corbett's argument, numerosity of a participant is more relevant to the conceptualization of the event than the numerosity of an event is relevant to the conceptualization of its participant. For discussion of other asymmetries and their motivations, see Corbett (2000, 8.8).

To sum up, we have seen how the empirical domain of verbal number may be delimited from agreement (Durie 1986) and from aspect, and how the verbal number itself is organized as two distinct domains, *participant number* vs. *event number* (Corbett 2000), the latter often also referred to as *pluractionality* (see Newman 2012, Mattiola 2020 for alternative terminological choices). However, these boundaries do not amount to categorical oppositions between mutually exclusive values, and we have seen that there are empirical reasons for allowing for Corbett's 'mixed' type of verbal number. Overall, while the conceptual distinctions are clear, the empirical details may be fuzzy in specific systems. Thus, the difference between verbal number and number agreement is clearer in languages with agreement which show accusative alignment, because verbal number tends to work on an ergative basis; and clearer in the languages which have different inventories of number values on nouns and on verbs than in the languages where these inventories are the same. In this volume, Krasnoukhova's comparative survey of South American systems has provided significant empirical evidence in support of the notion of a "mixed" verbal number (see her Fig. 2), and it is to be hoped that further cross-linguistic surveys may confirm or correct the main conclusions.

7 Number and conceptualization

There are empirical reasons to distinguish different dimensions of meaning in grammatical number, which correspond to distinct types of conceptual characterizations (see also Moravcsik 2017 for a detailed discussion, in particular on different conceptual perspectives underlying pluralization). Numerosity, based on the notion of cardinality and centred on the opposition between one and more than one, is the central one and it underpins the key singular-dual-plural contrast. A second dimension is related to the assessment of the size of the denotation. This is a scalar notion which is not based on cardinality; it supports intuitions about what counts as a small or a non-small amount, which underlies the distribution of the paucal, as well as of what counts as a quantity that exceeds expectations, characteristic of higher

plural readings. Thirdly, collective and distributive markers point to a dimension of conceptualization based on the degree of cohesion between the elements in the denotation. This is what distinguishes collections characterized as grouped together from others characterized as having members independent from each other.

These three dimensions are conceptual because they correspond to characterizations independent of language, rooted respectively in the ability to discriminate between different cardinalities, to assess relative magnitudes without counting, and to judge whether objects are related together by functional, spatial, or temporal proximity. They usefully organize the range of meanings that are expressed by number values, but they do not simply coincide with those meanings. Number values are language-internal categories, which can express a variety of semantic readings, not all straightforwardly describable in terms of those three conceptual dimensions. As we will see in this section, the details vary across grammatical systems, and the meaning of a given reading for a number value can integrate characterizations that belong to distinct conceptual dimensions, like being a small quantity and being cohesive.

The conceptual dimensions into which we want to organize the variety of attested interpretations are different from Haspelmath's (2010) *comparative concepts*. Like the notions that underlie the conceptual dimension we are considering, comparative concepts are defined on a language-independent basis and can serve as invariant reference categories to guide comparative analysis. However, Haspelmath explicitly characterizes comparative concepts as 'concepts created by comparative linguists for the specific purpose of cross-linguistic comparison.', without psychological reality. Our dimensions, on the other hand, are not such abstract constructs created to facilitate linguistic comparison. They are intended to be independently definable notions rooted not only in language but also in non-linguistic cognition. In the terminology advocated by Haspelmath, they are rather akin to 'conceptual-semantic concepts', except for the fact that they are not presented here as concepts but as conceptual dimensions along which the semantic content of number is articulated.

7.1 Numerosity and the singular-plural opposition

Consider first the most fundamental conceptual dimension, that of numerosity. The opposition between singularity and non-singularity is absolutely fundamental, but this does not mean that the grammatical opposition of singular and plural (or singular vs. all other values) coincides with it or is exhausted by it. We don't need to go beyond English (Germanic branch of Indo-European) to remind ourselves that the singular is also the default value used for mass nouns which do not define what counts as 'one' in their denotation domain, such as both concrete substance-denoting terms like *beer* and abstract notions like *fun*. Importantly, this includes terms

like *footwear* which are true of discrete individuals but describe them as a mass (in the extensive literature on them, see Barner and Snedeker 2005, Bale and Barner 2009, Chierchia 2010, Rothstein 2010, and also Corbett 2000, for a range of typological observations). Nor is it a novel observation that plurality can express not multiplicity, but manifold internal structure, as it often does in pluralia tantum like *suds* but also in pluralized mass nouns like *waters* (when they remain mass, not in their count reading as 'varieties / portions of water'; see Acquaviva 2008). Our sample confirms and strengthens this conclusion: the singular-plural opposition can certainly express aspects of mass interpretation, where numerosity is not relevant. For one thing, substance-denoting and abstract nouns which do not enter the opposition do not have to take singular as their only number value, but may be plural, like *mala* 'cold' or *madi* 'blood' carrying plural gender markers in Tswana, Southern Bantu (Creissels, this volume, 2.3.3). Likewise, mass nouns that have a fixed plural are reported to be "quite common" in Yimas, Lower Sepik (Foley, this volume, 548; the term for 'water' is formally a duale tantum). In Dargwa, East Caucasian, mass nouns are morphologically singular, and shift to 'kinds of' / 'portions of' interpretation when morphologically plural; but many of them control plural agreement even when morphologically singular; see a brief mention in Sumbatova (2018), and Corbett (2000: 261) on similar patterns in North America. In addition, a mass interpretation is compatible with a noun taking part in the singular-plural opposition, with pluralization often expressing abundance, as most clearly and productively in Arabic, Semitic (Fassi Fehri, this volume, 2.3), or heterogeneity, as for instance in Nivkh, isolate (Gruzdeva, this volume, 2.3.2.1, where an abundance reading is reported to have been available in previous stages of the language) or in Japonic languages (Shimoji, this volume, 2.5).

Another reason why the meaning of grammatical singularity and plurality is not exhausted by the opposition between 'one' and 'more than one' is that a collection may be linguistically characterized as 'one' too, if it is conceptualized as a bounded whole (see Moltmann 1997 for important discussion at the interface between philosophy and linguistic semantics). Aside from systems that specifically mark collective plurals, a collective reading is normally open to any plurally-marked NP (as in 'the boys lifted the piano [together]'), but the mutual cohesion which turns a plurality into a complex unit can also be expressed by the singular. In some East Caucasian languages, e.g. in Archi (East Caucasian), names of groups of people can be argued to show human plural agreement while being morphologically singular (Kibrik et al. 1977, and personal data). Arabic gives this a particularly clear grammatical expression (Fassi Fehri, this volume). Here, an underlying singular-(dual-)plural opposition interacts with a binary gender opposition (as well as with a collective-singulative opposition) to signal collective group readings by morphological means: these are feminine-marked singulars resulting from the affixation of forms that denote collections, and interpreted as collective wholes. In Fassi Fehri's own words, "a derived collective, termed plurative, denotes a group individual, or a plurality

perspectivized as a unity, rather than a multitude". The singular-plural opposition keeps its fundamental conceptual role, but the notion of "oneness" it translates into language-internal terms may refer to being a coherent whole, no matter how internally complex, rather than an undivided atom; Fassi Fehri aptly distinguishes the two senses as unit and atom.

A plural value may also express distinct readings which are related but conceptually quite distinct from each other. This happens when the marker which expresses the "simple" plural reading 'more than one N' on a noun N also expresses an associative reading for a suitable choice of N. An illustration from our sample is provided by Nganasan, Uralic (Gusev and Wagner-Nagy, this volume, 2.3.2.5), where plural morphology can also express 'individual N and associated people' and 'N and similar things' (this extends to the dual: 'N and one another individual associated with N'; see further examples of associative duals from various languages in Rukeyser 1997). Yet another example is what Corbett (2000: 241) terms anti-associative, where a set of elements associated with a referent does not include the referent itself, such as a plural of place name used as a means to refer to its inhabitants (Bagirokova et al., this volume; see also Daniel 2020a). These are all varieties of plurality (see Moravcsik 2017), and they are related by a common semantic denominator; but the different senses (criteria for belonging to the plural denotation) are not reducible to this common denominator. At most, we can say that the referents described by these heterogeneous senses all have a cardinality of more than one. This is the invariant component also present in pronominal plurals, otherwise very different from nominal pluralities.

In sum, the values of singular and plural are based on the conceptual opposition between one and more-than-one, but the meaning of elements marked by these values is not reducible to that.

7.2 Quantity size and "approximative" values

When a language has dedicated forms for a paucal reading, the meaning of this value correlates quite straightforwardly with the conceptual notion of a "small" collection, applicable to sets of cardinality small enough to allow for immediate grasp of its cardinality, without needing to count (this is the phenomenon of subitizing, well known from psychology and whose effects in language were discussed in Hurford 2001, Wiese 2004, Harbour 2014; see also Margolis 2020 for a proposal about innate representations of such small numerical quantities). The special status of small quantities is a fact about human (and animal) cognition; but the paucal is a linguistic category, and so its semantic range can differ from one language to another. Corbett (2000: 22–25) and Harbour (2014: 198–199) offer several examples, and in our sample, Foley for Yimas, Lower Sepik, sets the range at 'from three to about seven' for plural nouns in construction with a paucal-marked verb (not from two, since, in Yimas, that would call for the dual).

It is important to stress that the property of being a small collection is just as well-defined as the property of being a unit or a sum of units. What is undefined is the numerical cut-off point between what counts as a small collection and what is larger; but not being strictly defined in numerical terms does not make this property any less precise. Like being warm or tired, the property of being a collection of a size that qualifies it for the paucal value identifies a point along a scale not directly in terms of cardinality but on the basis of a more complex and partly context-dependent network of parameters. Harbour (2014: 199) made it clear that the numerically blurry boundaries of the paucal are an essential part of its content:

> The range of values that the paucal may cover within and across languages strongly suggests that it is correct not to rigidify the cardinal range as part of the inherent meaning of the feature, but to allow different speech communities to subject it to their own conventions.

As an empirical hypothesis, we may view the paucal as governed by the same sort of intuitions that underlie the use of determiners like *a couple* (also noted by Harbour, along with its German, Germanic, counterpart *ein paar,* as originally dual-meaning determiners which evolved towards a paucal reading, a development also attested, for instance, in Russian, Slavic; see also Plank 1989 on *duopaucal*). These intuitions are robust: speakers know whether a quantity qualifies as 'a couple' or not, even though this knowledge does not involve a cardinality. The lack of a clear numerical cut-off point makes these expressions of quantity akin to any other predicate that defines a point on a scale exceeding an unexpressed standard, like *large* or *cold* (in the positive grade; comparatives are different, see Kennedy and McNally 2005). The same applies to less frequently attested values like greater plurals (see Corbett 2000: 30 and Harbour 2014: 199 and 201 n13). These do not feature in our sample, except in the abundance reading discussed by Fassi Fehri in connection with Arabic "doubly pluralized" nouns (see above 4.4), and thus as a reading of plural but not as a number value individuated by a specific morphology. They must be mentioned in this connection, however, insofar as they, too, like paucals, are not defined in terms of cardinality but in terms of quantity size. The size they describe is also "approximative": not defined in terms of cardinality but rather as a quantity that exceeds expectations.

A final remark about the "approximate" semantic range of these values highlights a circumstance mentioned in the introduction to this section, namely the interconnection between conceptually distinct dimensions. Consider how Foley characterizes the semantic range of plural nouns in the Lower Sepik language Yimas (Foley, this volume, 548):

> The reading of a pluralized noun depends in Yimas on whether the verb governing it is marked for paucal number or not. If it is, the noun will be interpreted as referring to a set of objects, from three to seven or so – the upper bound not being rigid, though the bottom is, as paucal directly contrasts with dual, which is strictly 'two'. [...] The interpretation of plural in Yimas is normally a larger group, over seven, but again this is not rigid; the contrast between paucal

and plural is to highlight a smaller group versus a larger one, and so depending on context the boundary between them can be shifted up or down.

Apart from being related to the distinct conceptual dimensions of numerosity and quantity size, the plural and the paucal are distinct values in this system, applying to collections of distinct sizes despite the blurred boundary. This means that when a plural noun is in construction with a verb marked for paucal, the semantic range of the plural changes significantly, as it incorporates the range of another value in the paradigm. But "plurality greater than paucal-sized" and "plurality greater than two" are not the same notion. Stating the semantic range of the plural value in this language, then, requires listing both notions – which is exactly what Foley did.

7.3 Cohesion and distributive and collective

The contrast in cohesion, between collections characterized as grouped together or having members independent from each other, belongs to a third dimension of conceptualization, most often associated with verbal number and with the relation of participants, times, or location to the event or events. Being distributed or scattered is obviously closely related to being multiple, but it is not an indication of the cardinality or of the size of the collection. It comes as no surprise, then, that the grammatical expression relative to this parameter may co-occur with indications of numerosity, or may occur widely even though the language has otherwise quite limited expressions of number contrasts. Such is the case of Mohawk, Iroquoian, for example, notable among other things because the distributive clitics occurring on nominals denote multiple types, instead of simply denoting a plurality of tokens of the type described by the noun; see Mithun (this volume, 2.3.1).

The label 'distributive' applies to plurally-marked verbs (as in Karko, Nilo-Saharan; Marori, Trans-New Guinea; Nivkh, isolate; and various languages of South America) or to a core reading of reduplicated verbs or nouns (Indonesian, Malayo-Polynesian; Nivkh; Russian Sign Language; Mohawk, Iroquoian; Eastern Dan, Mande), but there are also several accounts of dedicated distributive verbal morphemes (Kiowa, Kiowa-Tanoan; Kakataibo, Panoan). Over and above the significant differences, these are all ways to signal that the event is scattered across distinct times, or locations, or participants, in a way that implies plurality as a necessary component of distinctness (see the discussion in section 6). Importantly, systems like Eastern Dan or Japonic languages show that the content of reduplication is distinct from the more semantically specialized marking of plurality, which is a plural word in Eastern Dan and a suffix in Japonic (Vydrin, this volume, 2.4.1, and Shimoji, this volume, 2.5).

The collective reading, which instead subsumes a plurality under a higher-level unity and likewise may apply to the event proper or its relationship with some arguments, is also ubiquitous as a possible interpretation of plural; we saw above how a

system like Arabic may specialize a feminine singular in this function (Fassi Fehri's 'plurative', mentioned in 7.1 above). What remains to consider now is the appearance of a dedicated collective morphology. In this connection we must make, or rather repeat, a terminological caveat. As Gil (this volume) reminds us, the term 'collective' "is used in a bewildering variety of ways to describe a wide range of linguistic phenomena". Even focusing on morphological forms specialized in this function, we must sharply distinguish forms expressing a cohesive group interpretation from 'collectives' understood as the semantically plural base forms opposed to derived, semantically singular singulatives. Fassi Fehri discusses how this intersects the singular-plural opposition in Arabic; see also Jakobi and Dimmendaal (this volume), for a discussion of Karko nouns that designate a collection in their base form and need a singulative affix to designate its one element. Only in the former understanding, not as the opposite of singulatives, does a collective specifically provide a grammatical expression for a cohesive interpretation. While comparatively rare, this is attested in our sample, namely by Gooniyandi, Bunuban, (dual and) plural enclitics, which regularly entail a collective reading, and by the collective affixes of Kakataibo, Panoan, which are similarly specialized but can only apply to nouns of ancestors or of animals, the latter only in tandem with the generic affix to denote species (McGregor, this volume, 2.1, Zariquiey, this volume, 2.2.2.4). The 'collectives' in the other sense, as opposed to singulatives, do not systematically entail a collective reading. However, there are cases where the terms 'collective' and 'singulative' refer not to a formal but a semantic opposition, between two readings of homophonous nouns; in these cases, the collective usually denotes granular substances (like 'snow' or 'firewood'), but it can also denote pluralities of things that make up a cohesive whole, like paired body parts in Nivkh (Gruzdeva, this volume, 2.3.2.2), or institutions like 'police' or 'army' in Indonesian (Gil, this volume, 2.3.2).

8 Conclusion and perspectives

Perhaps the most concise way to sum up the future avenues of research that this book suggests is "more of the same". It is in the nature of comparative enterprises to be open-ended, as any tentative conclusion calls for a larger empirical basis and raises further questions. But there are more specific reasons to wish for broader and deeper comparative studies based on something like our questionnaire (cf. Aikhenvald 2018 for an alternative).

First, we have acknowledged that our sample is not large enough to provide a reliable picture of the whole typological spectrum. In particular, we have been biased towards languages with rich number, so it would be instructive to supplement a larger selection of languages with systems where number has a very restricted grammatical expression. The closest to this in our sample is probably Indonesian, yet even Indonesian has extensive use of reduplication; more revealing comparisons

would emerge from systems such as South East Asian languages like Chinese (Sinitic). For similar reasons, it would be instructive to extend the investigation to more languages which problematize the very notion of word class, or lexical category. Indonesian (Malayo-Polynesian) and Mohawk (Iroquoian) in our sample have provided very useful illustrations of systems where the differences between nominal, pronominal, and verbal number must be approached as properties of stems in a morphosyntactic context, but clearly more such examples would be needed to ascertain what consequences "flexible" parts of speech have on the way a language encodes numerosity.

Regarding methodology, it would be interesting to extend the studies of the languages considered here to larger families (a development explicitly envisaged by Harbour and Mackenzie, this volume). This has to be done not so much to expand the empirical base, but more specifically because closely related (and so directly comparable) systems can bring into sharp focus the parameters of variation evidencing what tends to remain constant and how much other properties can change, and highlighting systemic links between parameters of variation, in line with the intra-genealogical typology (Greenberg 1969, Kibrik 1998, Šluinskij 2014). To some extent, our selection has included chapters framed in this comparative light (see especially the chapters on South American languages, contact languages, and the Lower Sepik family, and to a lesser extent the chapters on Nivkh, isolate, and Karko, Nilo-Saharan), but of course the room for further progress in this direction is vast. On a similar, but more practical side, the chapters may serve as leads for linguists working on description of (genealogically or areally) related languages.

There is one other line of inquiry that we envisage for the future, which emerges naturally from the structure of the questionnaire and from the use of the terminology deployed there and throughout the book. While there has not been any major terminological innovation, the attention to inter-comparability between individual studies has brought to the fore the need to make number-related terminology more uniform in a few areas. The use of the term *collective* in too many diverse senses has been rightly denounced as a practical hindrance (see comments in Gil, this volume), but it is convenient to keep that term in its relatively precise (if not formally defined) sense of "grammatical expression of plurality that makes up a cohesive unit", especially as the functional opposite of *distributed* (plural). The use of *collective* as opposed to singulatives, on the other hand, does not fall under this reading and is therefore best avoided, at least when it could give rise to confusion whether the non-singulative forms denote open-ended pluralities (as they usually do) or bounded units. Relatedly, Fassi Fehri distinguishes very usefully the two senses of 'one' by the terms *unit* and *atom*, and the conceptual distinction deserves to be firmly established (as well as the terminological one). In other cases we have chosen one term, or conversely one sense of a term, among several others adopted in the literature, simply to call for consistent and clear usage. We thus opt for *dyads* rather than natural pairs in the sense of cohesive sets of cardinality two (not to be con-

fused with *dyadic kinship terms*, see Evans 2006, also discussed in e.g. Bagirokova et al., this volume, Krasnoukhova, this volume, Gusev and Wagner-Nagy, this volume, McGregor, this volume); for *associative* plurals limited to nominal pluralities and more specifically human-based plurals meaning 'X and his/her group' rather than extended to, for instance, pronominal plurals, for which we suggest to use more generally *heterogeneous* plurals, as opposed to *additive* plurals (see also Moravcsik 2003). Contra Newman (2012) and Mattiola (2020), we argue against *pluractionality* as an alternative for *verbal number* because the former presupposes the numerosity of events and thus, for us, also distributivity, while the latter may simply indicate numerosity of participants, including when acting together. At the same time, as discussed in this chapter, we have avoided *clusivity* (together with *first person* inclusive and exclusive) and *optional number marking* because of our reservations about the validity of these concepts as analytic tools. Finally, we use *numerosity* as a label for the conceptual domain corresponding to the grammatical category of number as a whole, similarly to how tense is linguistic categorization of time (or temporal reference).

The format of our survey makes it a natural next step to expand the empirical base while ensuring comparability through a detailed questionnaire. This can be done both extensively by making the sample more representative by the inclusion of more languages from other families and areas and intensively with family/area surveys to account for micro-variation between related systems. Apart from the intrinsic interest of such an expansion, a development of our survey could also address some of the existing analytical questions, which the studies in this book help present in a clearer light.

The most obvious question is probably the one about the key parameters of variation. We know, and we have seen here, that large-scale schematic contrasts such as having "obligatory" or "non-obligatory" number marking, or being based on a singulative-collective opposition rather than singular-plural, or having or not having verbal number, are too coarse-grained. But even calling for finer-grained distinctions, while obviously justified, does not yet tell us what the real parameters of variation are, in terms of a predictive theory of how numerosity can and cannot be expressed in a grammatical system. Systematic comparative analyses are the necessary basis for any such theory, and the material gathered in this book represents a step in that direction. We have seen, for example, that "optionality" does not really mean that number marking – for any of its values – is truly a matter of unconstrained choice, but that it is (even if not straightforwardly or uniquely) associated with a certain interpretation: a nominal may be interpreted as plural, say, without this mandating overt marking, because such marking is not conditional on a plural reading per se, but on a plural reading that has particular informational prominence. It goes without saying that this is not an analysis, but an illustration of what "optional" might mean. When number marking is not systematically present or absent according to clear rules, as in Marori (Trans-New Guinea) derived nominaliza-

tions, in Arka and Dalrymple (this volume, 2.3), or on participles in Karko (Nilo-Saharan), in Jakobi and Dimmendaal (this volume, 2.3.1), what looks like a capricious distribution may well conceal deeper, previously unrecognized regularity; cf. the discussion of possible overt plural marking on interrogative pronouns in West Circassian, Northwest Caucasian (Bagirokova et al., this volume, 4.1), and especially the "sporadic" pattern of overt marking on agreeing adjectives and action nominals in Ket, Yeniseian (Vajda, this volume, 3.2). While it is already clear that languages vary a lot along this parameter, accounting for the logic behind the distribution of such "optional" number marking in specific languages requires more corpus or experiment-oriented research. This involves, among other things, making explicit the semantic and pragmatic conditions which license it; to put it differently, it involves finding out when number marking is not optional in a specific language.

The deepest set of questions concerns the essence of number as a linguistic category: not the patterning of number-related phenomena per se, but the primitives we want to posit, and the laws governing them, in order to account for the attested phenomenology. It is apparent that questions at this level necessarily presuppose fundamental theoretical choices. To contribute to this (necessarily ongoing) debate, we will conclude by identifying three key questions.

Firstly, the analysis of number variability between nouns, pronouns, and verbs turns on how one conceives of lexical categories. In most cases this looks like a non-issue, because the differences between them are self-evident. But systems where the same morphological root can display nominal or verbal morphosyntactic behaviour (Mohawk and West Circassian) or pronominal or nominal behaviour (see the discussion by Gil, this volume, on Indonesian) should remind us that membership in these categories may depend on the grammatical context, up to, in some approaches, leading to considering abolishing the distinction altogether. Regardless of whether or not parts of speech correspond to truly universal notions (see Croft 2000, Baker 2003, Haspelmath 2012, Rijkhoff and van Lier 2013, Panagiotidis 2014, Baker and Croft 2017, in the extensive literature), the attested differences between how number is organized across them call for an explanation; at least, we would like to be in a position to circumscribe the variation space. In fact, the very behaviour in terms of number marking is what distinguishes, for Gil (this volume), nouns and pronouns.

A recent contribution by Ghomeshi and Massam (2020) argues that pronominal and nominal number are in fact two distinct systems in the sense that at an appropriate level of analysis, they are properties of grammatical objects that differ in structure and not only in their categorial labelling; namely, they would respectively characterize the expression of person which defines pronouns, and the expression of part structure that is a sub-element of the structure which defines nouns.

From a different, functional perspective we have discussed the same dichotomy in section 5 above. The empirical success of such proposals, or of alternative ones, i.e. those treating (pro)nominal number in one piece, can be checked against what is known about the possible differences between nominal and pronominal number

systems. Ultimately, it may turn out that the degree of linguistic separation between the nominal and pronominal number is itself a parameter of cross-linguistic variation.

The second theoretical question foregrounded by the studies in this book concerns the bundling of number with other categories, such as person (as we have just seen in connection with pronouns), gender and noun class (Corbett 1995: 132–133, Aikhenvald 2014), and also probably less expectedly discourse-related categories like familiarity or definiteness (e.g. especially in connection with the dual, as mentioned in 4.6 above, but sometimes also with plural or 'collective', as in Gooniyandi, Bunuban, or Kakataibo, Panoan). Can number information be specifically associated with just any grammatically relevant category, or with any of its values? If only some types of pronouns are marked for number, for instance, can we predict what they will be, and why? Can any gender value be associated with any number-related reading, like feminine is associated with the unit reading in Arabic, or are there grammar-internal constraints? Is it accidental that the vast majority of mass nouns in Archi (East Caucasian) belong to the gender whose exponence is the same in the singular and in the plural (incidentally, zero)? Considering the contributions to this volume taken together underlines the need for a comparative investigation. In particular, it would be instructive to investigate the links of nominal number with information typically associated with nominal determiners, at phrasal level, like definiteness. This is an important aspect of the more general question of how number information is "distributed" within the noun phrase, an issue that has come to the fore in formal syntactic and semantic studies. Work in the former (see Landau 2016 for a concrete analysis which summarises various recent approaches) has highlighted the need to distinguish number as a morphological specification of a head noun from number as a morphosyntactic specification of the noun phrase, driving agreement with the verb. As for formal semantics, a significant development since Sauerland (2003) is the contention (far from universally accepted) that the primary syntactic locus for nominal number is *always* the Determiner head, and not just when plurality only arises at phrasal level like in conjoined nouns (the phrase [*Bonnie and Clyde*] is plural even though neither of its parts is). In this formal theoretical context, a clearer picture of the links between N and D (and/or between number on N and D-related information) in an empirical typological perspective would evidently prove very useful. See also Corbett (1991, Chapter 9; 2006: 168–170) for a functional discussion of related phenomena.

Thirdly and finally, any attempt at explaining why languages express number properties in the way they do must be based on a view of what exactly we understand as the content of the category of number. This means not just its non-linguistic conceptual basis, but also what kind of information is expressed by the semantics of number as part of language structure, what are its basic building blocks, how universal they are, and how they can and, especially, how they cannot combine. These are of course very abstract questions, but to see their relevance, consider that

some definition of what we mean by "number" is needed before making any statement about languages where this category is completely absent (Corbett 2000: 50–51 tentatively mentions Pirahã, isolate; Old Javanese, Malayo-Polynesian branch of Austronesian; and Old Chinese, Sinitic, as possible examples). It is an empirical fact that languages do not need to encapsulate number-related information in their nominal and verbal morphology; but whether a language may dispense completely with all number-related information clearly depends on what one understands by such information. The point is not merely terminological because one's understanding of number affects what one thinks about the boundaries of number, about any supposedly "impossible" system, and about the reasons, systemic, diachronic or systemic-diachronic, for the frequency and rarity of certain patterns. One approach that has emerged in the literature addresses this concern about predictivity by proposing that the attested values are in fact epiphenomena; notions like "singular", "unit augmented", or "paucal" arise from the combination of a small number of basic features (semantic specifications which restrict an expression's denotation), and only these features have cognitive reality. Work in this line (see Noyer 1997, Nevins 2011, Harbour 2011, Martí 2020) has arguably culminated in the comprehensive proposal articulated in precise formal terms by Harbour (2014). While this specific approach has been defended by researchers working in formal frameworks, this way of thinking is not incompatible with typological investigative enterprises (see e.g. Cysouw 2001), and the proposals in fact explicitly set out to predict the constraints on observed number systems, represented most clearly by the implicational universals established by Greenberg's (1963) universals 34 (a trial presupposes a dual and a dual presupposes a plural) and 35 (plural, unlike singular, always has non-zero allomorphs) within the purely typological tradition. But any success towards this goal is predicated on a detailed and comprehensive typology of number systems; clearly, it is necessary to ascertain the empirically attested patterns in order to account for them. This book, as stated in chapter 1, aims to contribute towards this goal, with the wealth of data and analyses detailed in the individual chapters, and along the broad themes summarized in this concluding chapter.

Acknowledgments

References

Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2003. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2014. Number and noun categorisation. A view from north-west Amazonia. In Anne Storch & Gerrit J. Dimmendaal (eds.), *Number – Constructions and Semantics. Case studies from Africa, Amazonia, India and Oceania*, 33–55. Amsterdam: Benjamins.

Aikhenvald, Alexandra Y. 2018. Number systems in grammar. Position paper. Available as https://www.jcu.edu.au/__data/assets/pdf_file/0004/789700/Number-systems-in-grammar-Position-paper.pdf

Allan, Keith. 1980. Nouns and countability. *Language* 56. 541–567.

Andrzejewski, Bogumił W. 1960. The Categories of Number in Noun forms in the Borana Dialect of Galla. *Africa* 30(1). 62–75. doi:10.2307/1157743.

Austerlitz, Robert. 1959. Semantic Components of Pronoun: Gilyak. *Word* 15. 1, 102–109, DOI: 10.1080/00437956.1959.11659686

Baker, Mark. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

Baker, Mark & William Croft. 2017. Lexical categories: legacy, lacuna, and opportunity for functionalists and formalists. *Annual Review of Linguistics* 3. 179–197.

Bale, Alan & David Barner. 2009. *The Interpretation of Functional Heads: Using Comparatives to Explore the Mass/Count Distinction. Journal of Semantics* 26. 217–152.

Bale, Alan & Jessica Coon. 2014. Classifiers are for numerals not for nouns: Consequences for the mass-count distinction. *Linguistic Inquiry* 45. 695–707. doi: doi.org/10/1162/LING_a_00170

Barner, David & Jesse Snedeker. 2005. Quantity judgments and individuation: Evidence that mass nouns count. Cognition 97. 41–66.

Barulin, Aleksandr N. 1980. Kategorija čisla v mestoimenijax [Number in pronouns] In Vladislav M. Andrjuščenko (ed.), *Issledovanija v oblasti grammatiki i tipologii jazykov* [Investigations in grammar and typology of languages], 145–163. Moscow: MGU.

Bhat, D. N. S. 2004. *Pronouns*. Oxford: Oxford University Press.

Birjulin, Leonid A. 1989. Kvantifikacija predikatov. In Viktor S. Xrakovskij (ed.), *Tipologija iterativnyx konstrukcij*, 243–258. Leningrad: Nauka. [English revised translation published as Xrakovskij, Victor. 1997. "Typology of iterative constructions." Muenchen: Lincom Europa].

Bisang, Walter. 2010. Word classes. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology,* 281–300. Oxford: Oxford University Press.

Bloomfield, Leonard. 1957. *Eastern Ojibwa: Grammatical sketch, texts, and word list*. University of Michigan Press.

Borer, Hagit. 2005. *In Name Only*. Vol. 1 of *Structuring Sense*. Oxford: Oxford University Press.

Boxwell, Maurice. 1967. Weri Pronoun System. *Linguistics* 5(29). 34–43.

Brugmann, Karl. 1900. *Griechische Grammatik*. München: Beck.

Bylinina, Lisa & Alexander Podobryaev. 2020. Plurality in Buriat and Structurally Constrained Alternatives, *Journal of Semantics* 37(1). 117–128. https://doi.org/10.1093/jos/ffz017

Cabredo Hofherr, Patricia & Brenda Laca (eds.). 2012. *Verbal plurality and distributivity*. (Linguistische Arbeiten, Vol. 546). Berlin/Boston: Walter de Gruyter.

Chierchia, Gennaro. 1998. Plurality of Mass Nouns and the Notion of Semantic Parameter. In Susan Rothstein (ed.), *Events and Grammar,* 53–103. Dordrecht: Kluwer.

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149.

Conklin, Harold C. 1962. Lexicographical treatment of folk taxonomies. In F. W. Householder & Sol Saporta (eds.), *Problems in lexicography. International Journal of American Linguistics* 28(4). 119–141.

Corbett, Greville G. & Richard J. Hayward. 1987. Gender and number in Bayso. *Lingua* 73(1–2). 1–28.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2005. Suppletion in personal pronouns: Theory versus practice, and the place of reproducibility in typology. *Linguistic Typology* 9. 1–23.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2010. Implicational Hierarchies. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.

Croft, William. 2000. Parts of speech as typological universals and as language particular categories. In Petra Maria Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 65–102. Berlin: Mouton de Gruyter.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Cusic, David D. 1981. *Verbal plurality and aspect*. Diss. Stanford University.

Cysouw, Michael. 2001. *The paradigmatic structure of person marking*. [Reprinted as Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press].

Cysouw, Michael. 2005. Syncretisms involving clusivity. In Elena Filimonova (ed.), *Clusivity: Typology and case studies of inclusive-exclusive distinction,* 73–111. (Typological Studies in Language 63). Amsterdam: John Benjamins Publishing.

Dahl, Osten & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorsten Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility,* 47–64. Amsterdam: Benjamins.

Dalle Ceste, Carlo. Ms. Three is a company. The emergence of number-values in Western Oceanic pronouns. Under revision for *Studies in Language*.

Daniel, Michael. 2005. Understanding inclusives. In Elena Filimonova (ed.), *Clusivity: Typology and case studies of inclusive-exclusive distinction*, 3–48. Amsterdam: John Benjamins Publishing. doi:10.1075/tsl.63.04dan.

Daniel, Michael. 2013. Plurality in Independent Personal Pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/35, Accessed on 2021 08 05.)

Daniel, Michael. 2020a. Associative plural as indexical category. *Language Sciences* 81. https://doi.org/10.1016/j.langsci.2019.101256

Daniel, Michael. 2020b. Person in Morphology. *Oxford Research Encyclopedia of Linguistics*. doi:10.1093/acrefore/978019938465

Daniel, Michael & Edith Moravcsik. 2013. The Associative Plural. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/36, Accessed on 2021 08 05.)

Dimmendaal, Gerrit. 2000. Number marking and noun categorization. *Anthropological Linguistics* 42/2. 214–261.

Dixon, Robert. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Doetjes, Jenny. 2012. Count/mass distinctions across languages. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, Volume 33.3, 2559–2580. Berlin: De Gruyter Mouton.

Dolinina, Inga B. 1989. Teoretičeskie aspekty glagoľnoj množestvennosti. In Viktor S. Xrakovskij, (ed.), *Tipologija iterativnyx konstrukcij*, 258–269. Leningrad: Nauka. [English revised translation published as Xrakovskij, Victor. (ed.). 1997. *Typology of iterative constructions*. Muenchen: Lincom Europa].

Dolinina, Inga B. 1990. Glagoľnaja množestvennosť: soderžateľnye tipy [Verbal plurality: a functional classification]. In Leonid A. Birjulin, Igor' F. Varduľ & Viktor S. Xrakovskij (eds.), *Tipologija i grammatika* [Typology and grammar], 141–151. Moscow: Nauka.

Dressler, Wolfgang U. 1968. *Studien zur verbalen Pluralität: Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen*. (Sitzungsberichte der Österreichischen Akademie der Wissenschaften 259/1). Wien: Böhlau.

Dryer, Matthew. 1989. Plural words. *Linguistics* 27(5). 865–896.

Durie, Mark. 1986. The Grammaticization of Number as a Verbal Category. *Annual Meeting of the Berkeley Linguistics Society* 12. 355–368. doi:10.3765/bls.v12i0.1876.

Evans, Nicholas. 2006. Dyadic constructions. In Keith Brown (ed.), *Encyclopedia of language and linguistics,* 24–28. Oxford: Elsevier.

Filimonova, Elena (ed.). 2005. *Clusivity: Typology and case studies of inclusive-exclusive distinction*. (Typological Studies in Language 63). Amsterdam: John Benjamins Publishing.

Fillmore, Charles J. 1968. The Case for Case. In Emmon Bach & Robert T. Harms (eds.), *Universals in Linguistic Theory*, 1–88. New York: Holt, Rinehart, and Winston.

Forchheimer, Paul. 1955. *The category of person in language*. Berlin: De Gruyter.

Frajzyngier, Zygmunt. 1985. Ergativity, number, and agreement. *Annual Meeting of the Berkeley Linguistics Society*. Vol. 11. 96–106.

Ghomeshi, Jila & Diane Massam. 2020. Number is different in nominal and pronominal phrases. *Linguistic Inquiry* 51. 597–610.

Gil, David. 1993. Nominal and verbal quantification. *STUF-Language Typology and Universals* 46, no. 1–4. 275–317.

Gil, David. 2013. Numeral Classifiers. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/55, Accessed on 2021 08 05.)

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language,* 73–113. Cambridge, MA: Cambridge University Press.

Greenberg, Joseph H. 1969. Some methods of dynamic comparison in linguistics. In Jaan Puhvel (ed.), *Substance and structure of language*, 147–203. Berkeley & Los Angeles: University of California Press.

Greenberg, Joseph H. 1988. The First Person Inclusive Dual as an Ambiguous Category. *Studies in Language* 12(1). 1–18. doi:10.1075/sl.12.1.02gre.

Greenberg, Joseph H. 1989. On a metalanguage for pronominal systems: a reply to McGregor. *Studies in Language* 13(2). 437–451.

Gruzdeva, Ekaterina. 1998. *Nivkh*. München: LINCOM Europa.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2021. *Glottolog* 4.4. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.4761960. (Available online at http://glottolog.org, Accessed on 2021 09 01.)

Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42. 561–594.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90. 185–229.

Harbour, Daniel. 2016. *Impossible persons*. The MIT Press.

Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86. 663–687.

Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything: In honor of Edward Keenan*, 109–130. Los Angeles: UCLA Working Papers in Linguistics.

Haspelmath, Martin. 2013. Occurrence of Nominal Plurality. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck

Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/34, Accessed on 2021 08 05.)

Hurford, James. 2001. Languages Treat 1–4 Specially. *Mind and Language* 1. 69–75.

Iljic, Robert & Luo Yi（罗毅）. 2001. The problem of the suffix -men in Chinese grammar. *Journal of Chinese Linguistics* 29(1). 11–68.

Ingram, David. 1978. Personal pronoun typology and universals. In Joseph Greenberg (ed.), *Universals of human language,* Vol. 3. 213–247.

Ivani, Jessica Katiuscia. 2017. *The morphosyntax of number systems: a cross-linguistic study.* Università degli Studi di Bergamo / Università degli Studi di Pavia Dissertation.

Ivani, Jessica & Taras Zakharko. 2019. *Tymber, Database on nominal number marking constructions.* Version 1.0.0.-alpha. https://github.com/jivani/tymber/

Jespersen, Otto. 1924. *The philosophy of grammar.* Allen & Unwin.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree mofification, and the semantics of gradable predicates. *Language* 81. 345–381.

Kibrik, Aleksandr E., Sandro V. Kodzasov & Irina P. Olovjannikova. 1972. *Fragmenty grammatiki xinalugskogo jazyka.* [Elements of the Grammar of Khinalug] Moscow: MGU.

Kibrik, Aleksandr E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka* [A structural description of Archi] Moscow: MGU.

Kibrik, Aleksandr E. 1998. Does intragenetic typology make sense? In Winfried Boeder, Christoph Schroeder, Karl-Heinz Wagner & Wolfgang Wildgen (eds.), *Sprache im Raum und Zeit: In memoriam Johannes Bechert: Beiträge zur empirischen Sprachwissenschaft,* 61–68. Tübingen: Narr.

Kibrik, Andrej A. 2011. *Reference in discourse.* Oxford: Oxford University Press.

Kim, Ronald I. 2018. *The Dual in Tocharian From Typology to Auslautgesetz.* Dettelbach: Verlag J.H. Röll GmbH.

Kiparsky, Paul. 2013. On the interaction and interpretation of nominal features. A Paper presented at Stanford University. Oct. 25, 2013.

Kouneli, Maria. 2020. Number-based noun classification: The view from Kipsigis. *Natural language and Linguistic Theory* 38. 1–57. doi.org/10.1007/s11049-020-09494-8

Krasnoukhova, Olga. 2012. *The noun phrase in the languages of South America.* Netherlands Graduate School of Linguistics. Utrecht: LOT.

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of Chinese and English. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The Generic Book,* 398–411. Chicago and London: The University Of Chicago Press.

Lamberterie, Charles de. 1979. Le signe du pluriel en arménien classique. *Bulletin de la Société de Linguistique de Paris Paris* 74(1). 319–332.

Landau, Idan. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language and Linguistic Theory* 34. 975–1020.

Lichtenberk, Frantisek. 2000. Inclusory pronominals. *Oceanic linguistics* 39(1). 1–32.

Malchukov, Andrej. 2013. Verbalization and insubordination in Siberian languages. In Martine Irma Robbeets & Hubert Cuyckens, *Shared Grammaticalization. With special focus on the Transeurasian languages,* 177–208. Amsterdam: John Benjamins.

Margolis, Eric. 2020. The small number system. *Philosophy of Science* 87. 113–134.

Martí, Luisa. 2020. Inclusive Plurals and the Theory of Number. *Linguistic Inquiry* 51. 37–74.

Massam, Diane. 2009. Noun incorporation: essentials and extensions. *Language and Linguistics Compass* 3(4). 1076–1096, doi:10.1111/j.1749–818x.2009.00140.x.

Massam, Diane (ed.). 2012. *Count and Mass across Languages.* Oxford: Oxford University Press.

Matasović, Ranko. 2019. A universal about adnominal number agreement. *Italian Journal of Linguistics* 31(2). 169–200.

Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A comprehensive grammar.* London: Routledge.

Mattiola, Simone. 2019. *Typology of pluractional constructions in the languages of the world*. Amsterdam: John Benjamins Publishing Company.

Mattiola, Simone. Pluractionality: A cross-linguistic perspective. *Language and Linguistics Compass* 14(3). doi:10.1111/lnc3.12366

Mauri, Caterina & Andrea Sansò. 2022. Heterogeneous sets: a diachronic typology of associative and similative plurals. *Linguistic Typology* 26(1). doi:10.1515/lingty-2021-2072.

McGregor, William B. 1989. Greenberg on the first person inclusive dual: Evidence from some Australian languages. *Studies in Language* 13(2). 452–458.

McGregor, William B. 1996. The pronominal system of Gooniyandi and Bunuba. In William B. McGregor (ed.), *Studies in Kimberley languages in honour of Howard Coate,* 159–173. München & Newcastle: Lincom Europa.

McKaughan, Howard P. 1959. Semantic Components of Pronoun Systems: Maranao. *Word* 15. 1, 101–102.

McKay, Graham R. 1978. Pronominal Person and Number Categories in Rembarrnga and Djeebbana. *Oceanic Linguistics* 17(1). 27–37.

McKay, Graham R. 1990. The Addressee: or Is the second person singular? *Studies in Language* 14(2). 429–432.

Meillet, Antoine. 1934. *Le Slave Commun*. Paris: H. Champion.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In Michael Hammond & Michael Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics,* 211–234. New York: Academic Press.

Mithun, Marianne. 1999. *The languages of native North America*. Cambridge: Cambridge University Press.

Moltmann, Friederike. 1997. *Parts and Wholes in Semantics*. Oxford: Oxford University Press.

Moravcsik, Edith. 1978. Agreement. In Joseph H. Greenberg (ed.) *Universals of Human Language 4: Syntax*, 331–374. Stanford: Stanford University Press.

Moravcsik, Edith. 1994. *Group plural – associative plural or cohort plural*. Email document, LINGUIST List, *5*.681. 1068–4875. (https://linguistlist.org/issues/5/5-681/)

Moravcsik, Edith. 2003. A semantic analysis of associative plurals. *Studies in Language* 27(3). 469–503. doi:10.1075/sl.27.3.02mor.

Moravcsik, Edith. 2017. Number. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *The Cambridge Handbook of Linguistic Typology,* 440–476. Cambridge: Cambridge University Press. doi:10.1017/9781316135716.014.

Nevins, Andrew. 2007. The representation of third person and its consequences for the person-case constraint. *Natural Language and Linguistic Theory* 25. 273–313. doi: dx.doi.org/10.1007/s11049-00609017-2

Nevins, Andrew. 2011. Marked Triggers vs. Marked Targets and Impoverishment of the Dual. *Linguistic Inquiry* 42. 413–444.

Nevins, Andrew & Mário Coelho da Silva. 2020. Maxakalí has suppletion, numerals and associatives but no plurals. *Linguistic Variation* 20(2). 27–287. doi: https://doi.org/10.1075/lv.00020.nev

Newman, Paul. 1980. *The Classification of Chadic within Afroasiatic*. Leiden: Universitaire Pers.

Newman, Paul. 2012. Pluractional Verbs: An Overview. In Patricia Cabredo Hofherr & Brenda Laca (eds.), *Verbal Plurality and Distributivity*. Berlin/ Boston: De Gruyter. 185–210. https://doi.org/10.1515/9783110293500.185

Nikolaeva, Irina. 2014. *A grammar of Tundra Nenets*. Mouton: De Gruyter.

Noyer, Rolf. 1997. *Features, Positions, and Affixes in Autonomous Morphological Structure*. New York: Garland.

Panagiotidis, Phoevos. 2014. *Categorial Features: A Generative Theory of Word Class Categories*. Cambridge: Cambridge University Press.

Panfilov, Vladimir Z. 1962. *Grammatika nivxskogo jazyka* [Grammar of Nivkh]. Part 1: Fonetičeskie vvedenie i morfologija imennyx častej reči [Introduction to phonetics and nominal morphology]. Moscow-Leningrad: AN SSSR.

Pertsova, Katya. (Ms.) A case for a binary feature underlying clusivity. Forthcoming in *Morphology*.

Plank, Frans. 1985. Die Ordnung der Personen. *Folia Linguistica* 29(1–2). 111–176.

Plank, Frans. 1989. On Humboldt on the Dual. In Roberta Corrigan, Fred Eckman & Michael Noonan (eds.), *Linguistic Categorization Current Issues in Linguistic Theory*, 293–334. Amsterdam/Philadelphia: Benjamins. doi:10.1075/cilt.61.18pla.

Quine, Willard van Orman. 1960. *Word and Object*. Cambridge, MA: MIT Press.

Rijkhoff, Jan & Eva van Lier. 2013. *Flexible Word Classes: Typological Studies of Underspecified*. Oxford: Oxford University Press.

Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27. 343–397.

Rukeyser, Alison Smiley. 1997. *A Typology of the Nominal Dual: Evidence from Indo-European, Finno-Ugric, Semitic and Australian Languages*. University of California, Davis.

Sato, Hiroko. 2013. *Grammar of Kove: an Austronesian language of the West New Britain province, Papua New Guinea*. University of Hawai'i at Mānoa.

Sauerland, Ulrich. 2003. A new semantics for number. In Rob Young & Yuping Zou (eds.), *SALT XIII,* 258–275. Ithaca, New York: CLC Publications.

Schwarzwald, Ora. 1991. Grammatical vs. lexical plural formation in Hebrew. *Folia Linguistica* 25. 577–608.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Simon, Horst J. 2005. Only *you?* Philological investigation into the alleged inclusive-exclusive distinction in the second-person plural. In Elena Filimonova (ed.), *Clusivity: Typology and case studies of inclusive-exclusive distinction*, 113–150. (Typological Studies in Language 63). Amsterdam: John Benjamins Publishing.

Šluinskij, Andrej B. 2006. K tipologii predikatnoj množestvennosti: organizacija semantičeskoj zony. [Towards a typology of predicative plurality] *Voprosy jazykoznanija* 1 (2006). 46–75.

Šluinskij, Andrej B. 2014. Vnutrigenetičeskaja tipologija jazykov kva: morfosintaksis glagoľnoj gruppy [Intragenealogical typology of Kwa: Morphosyntax of the verbal phrase] In Viktor A. Vinogradov (ed.), *Osnovy afrikanskogo jazykoznanija. Diaxroničeskie processy i genetičeskie otnošenija jazykov Afriki,* 354–455. Moscow: Jazyki slavjanskoj kuľtury.

Smith-Stark, Thomas C. 1974. The plurality split. *Chicago Linguistic Society* 10(1). 657–672.

Sokolovskaja, Nataľja K. 1980. Nekotorye semantičeskie universalii v sisteme ličnyx mestoimenij [Some semantic universals in the system of personal pronouns]. In Igor' F. Varduľ (ed.), *Teorija i tipologija mestoimenij* [Theory and typology of pronouns]. Moscow: Nauka. 84–103.

Sumbatova, Nina R. 2018. Osnovanija imennyx klassifikacij: ot semantiki do fonologii [Principles of nominal classifications: from semantics to phonology]. *Voprosy jazykoznanija* 6 (2018). 7–30.

Stifter, David. 2013. Vocative for nominative. In Barbara Sonnenhauser & Patrizia Noel Aziz Hanna, (eds.), *Vocative! Addressing between System and Performance*, 43–85. (Trends in Linguistics. Studies and Monographs [TiLSM] 261). Berlin/Boston: De Gruyter Mouton.

Suprun, Adam E. & Aleksandr M. Moldovan. 2005. Staroslavjanskij i cerkovnoslavjanskij jazyk [Old Church Slavonic and Church Slavonic]. In Aleksandr M. Moldovan, Sergej S. Skorvid, Andrej A. Kibrik, Nataľja V. Rogova, Ekaterina I. Jakuškina & Svetlana M. Tolstaja (eds.), *Jazyki mira*. *Slavjanskie jazyki* [Languages of the World. Slavic Languages], 29–69. Saint Petersburg: Nestor-Istorija.

Tang, Marc & One-Soon Her. 2019. Insights on the Greenberg-Sanches-Slobin generalization: Quantitative typological data on classifiers and plural markers. *Folia Linguistica* 53. 297–331. doi.org/10.1515/flin-2019–2013

Tereščenko, Natalʹja M. 1965. *Nenecko-russkij slovar'*. [Nenets Russian Dictionary]. Moscow: Sovetskaja Ènciklopedija.

Thomas, David. 1955. Three Analyses of the Ilocano Pronoun System. *Word* 11(2). 204–208. doi:10.1080/00437956.1955.116595.

Tsiakmakis, Evripidis, Joan Borràs-Comes & Maria Teresa Espinal. 2021. The interpretation of plural mass nouns in Greek. *Journal of Pragmatics* 181. 209–226. doi.org/10.1016/j.pragma.2021.05.017.

Vaillant, André. 1958. *Grammaire comparée des langues slaves*. Tome II. Morphologie. Première Partie: Flexion nominal. Lyon: IAC.

Vassilieva, Maria Borisovna. 2005. Associative and Pronominal Plurality. Ph. D. thesis. Stony Brook University.

Veselinova, Ljuba. 2006. *Suppletion in verb paradigms*. (Typological Studies in Language 67). Amsterdam: Benjamins Publisher.

Von Humboldt, Wilhelm. 1827. Über den Dualis. *Abhandlungen der historisch-philologischen Klasse der königlichen Akademie der Wissenschaften zu Berlin aus dem Jahre 1827*. Berlin: Druckerei der königlichen Akademie der Wissenschaften.

Von Soden, Wolfram. 1969. *Grundriss der akkadischen Grammatik*. Rome: Pontificium Institutum Biblicum.

Wiese, Heike. 2004. *Numbers, Language, and the Human Mind*. Cambridge: Cambridge University Press.

Xrakovskij, Viktor. S. 1987. Kratnostʹ. In Lia V. Bondarko (ed.), *Teorija funkcionalʹnoj grammatiki: Vvedenie. Aspektualʹnostʹ. Vremennaja lokalizovannostʹ. Taksis* [The theory of functional grammar. Introduction. Aspect. Temporal reference. Taxis]*, 124–152. Leningrad: Nauka.

Xrakovskij, Viktor S. 1989. Semantičeskie tipy množestva situacij i ix estestvennaja klassifikacija [Semantic types of event plurality and their natural classification]. In Viktor S. Xrakovskij (ed.), Tipologija iterativnyx konstrukcij. 5–53. Leningrad: Nauka. [English revised translation published as Victor Xrakovskij. 1997. Typology of iterative constructions. München: Lincom Europa].

Zhang, Niina. 2013. *Classifier Structures in Mandarin Chinese*. Berlin/Boston: De Gruyter Mouton. doi: 10.1515/9783110304992.

Zúñiga, Fernando. 2006. *Deixis and alignment: inverse systems in indigenous languages of the Americas*. Amsterdam: John Benjamins Publishing Company.

Zwicky, Arnold M. 1977. Hierarchies of person. In Woodfrod A. Beach, Samuel E. Fox & Shulamith Philosoph (eds.), *Papers from the Thirteenth Regional Meeting,* 714–733. Chicago: Chicago Linguistic Society.

Terms

Languages and language families

Authors