NULL SUBJECTS IN SLAVIC AND FINNO-UGRIC

LICENSING, STRUCTURE AND TYPOLOGY

Edited by Gréte Dalmi, Egor Tsedryk, Piotr Cegłowski

Null Subjects in Slavic and Finno-Ugric

Studies in Generative Grammar

Editors
Norbert Corver
Harry van der Hulst

Founding editors
Jan Koster
Henk van Riemsdijk

Volume 145

Null Subjects in Slavic and Finno-Ugric

Licensing, Structure and Typology

Edited by
Gréte Dalmi
Egor Tsedryk
Piotr Cegłowski

DE GRUYTER
MOUTON

Contents

List of contributors

Anders Holmberg, Professor, Department of Linguistics, University of Newcastle, Newcastle, UK

Piotr Cegłowski, Professor, Adam Mickiewicz University, Poznań, Poland

Gréte Dalmi, Professor, External Consultant, Hamburg University, Dept. of Finno-Ugric Studies, Hamburg, Germany

Nerea Madariaga, Professor, Department of Slavic Studies, University of the Basque Country, Vitoria-Gasteiz, Spain

Egor Tsedryk, Professor, Department of Languages and Cultures, Saint Mary's University, Halifax, Canada

Axel Holvoet, Professor, Institute for the Languages and Cultures of the Baltic, Vilnius University, Vilnius, Lithuania

Anna Daugavet, Researcher, Institute for the Languages and Cultures of the Baltic, Vilnius University, Vilnius, Lithuania

Beáta Wagner Nagy, Professor, Department of Finno-Ugric Studies. University of Hamburg, Hamburg, Germany

Susann Fischer, Professor, Department of Romance Studies. University of Hamburg, Hamburg, Germany

Dobrinka Genevska-Hanke Assistant Professor, School of English and American Studies, Carl von Ossietzky University, Oldenburg, Germany

Jeremy Bradley, Researcher, Department of Finno-Ugric Studies, University of Vienna, Vienna, Austria

Johannes Hirvonen, PhD student, Department of Finno-Ugric Studies, University of Vienna, Vienna, Austria

Urpo Nikanne, Professor, Department of Finnish Åbo Akademi University, Turku, Finland

Marta Ruda, Associate Professor, Department of English Linguistics, Jagellonian University, Cracow, Poland

Ludmila Veselovská, Professor, Department of English Linguistics, Palacký University, Olomouc, Czech Republic

Mikael Vinka, Professor, Department of Saami Studies, University of Umeå, Umeå, Sweden

Jacek Witkoś, Professor, Department of Comparative English-Polish Studies, Adam Mickiewicz University, Poznań, Poland

List of abbreviations

ABL	ablative case
ACC	accusative case
ACT	active voice
ADE	adessive case ('at')
ADJ	adjectivizer
ADV	adverbial
AGRP	agreement phrase
AGROP	object agreement phrase
AGRSP	subject agreement phrase
ANIM	animate noun
ANT	anterior tense
AOR	aorist tense
ARB	arbitrary 'people' interpretation
ASP	aspect
ASPP	aspect phrase
A-Topic	aboutness shift topic
AU	auditive (reported) evidentiality, 'hearsay'
AUX	auxiliary verb
BRU	Belarusian, Russian, Ukrainian (East Slavic languages)
CAR	caritive suffix
CAUS	causative suffix
CG	common ground
CL	clitic auxiliary or pronoun
CO	co-affix
COM	comitative case ('with')
COND	conditional mood
ContrP	contrast phrase
CNT	contemporaneous tense
CNG	connegative form
CP	complementizer phrase
CS	causative situation
CVB	converb
DA	direct argument
DAT	dative case
DEB	debitive mood (obligation)
[±DEF]	[±definiteness] feature
DEM	demonstrative
DET	determiner
DIM	diminutive suffix
DIST	distal feature
DM	distributed morphology
DP	determiner phrase
DU	dual number
DUR	durative aktionsart

DTR	detransitivizing suffix
DST	destinative case
ELA	elative case 'out of'
EMPH	emphatic particle
EP	epenthetic vowel
EPP	Extended Projection Principle
ESS	essive case
EXCL	exclamative particle
F	feminine gender
FA	functional application
FamP	familiar topic phrase
FinP	finiteness phrase
FocP	focus phrase
FORM	formalis case
FRQ	frequentative aktionsart
FUT	future tense
GEN	genitive case
GF	grammatical function
GN	generic operator/feature
GP	ground phrase
G-Topic	given/familiar topic
HAB	habitual aspect
HORT	hortative mood (encouragement)
i	speaker ('I')
ILL	illative case ('into')
IMP	imperative mood
INC	inchoative
INDF	indefinite
INE	inessive case ('in')
INF	infinitive
INFER	inferential mood
INS	instrumental case ('with/by')
INT	interrogative marker
INTR	intransitive verb
IPFV	imperfective aspect
IRR	irrealis mood
ITER	iterative aktionsart
LAT	lative case
LF	logical form
LOC	locative case
LOG	logophoric pronoun
M	masculine gender
N	neuter gender
NA	non-agreeing form
NEG	negation, negative operator/particle
NMLZ	nominalizer
NOM	nominative case

NP	noun phrase
nP	"little n" phrase
NS	null subject
NSL	Null Subject Language
o	other (neither the speaker nor the addressee)
OBL	oblique case/argument
OBJ	object
PA	predicate abstraction
PART	partitive case
PASS	passive voice/morpheme
PFV	perfective aspect
PFX	prefix
PL	plural number
PHON	phonological feature
PF	phonological form
PIC	Phase Impenetrability Condition
POSS	possessive marker/suffix
POST	posterior tense
POT	potentialis mood 'possibility'
PP	prepositional phrase
PRF	perfect aspect
pro	phonologically null pronominal in finite clauses
pro-drop	pronoun drop
PROX	proximate feature
PRS	present tense
PST	past tense
PTCP	participle
PTCP$_A$	active participle
PTCP$_P$	passive participle
QP	quantifier phrase
REL	relative pronoun
RFL	reflexive
RES	resultative suffix
RPO	reflexive possessive
S	set
SAPP	speech act participant phrase
SBJV	subjunctive mood
SHIFTP	shift topic phrase
SG	singular number
SPEC	specifier
SOT	sequence of tenses
SPR	superessive case ('on top of')
SUBJ	subject
SUP	supine
t	trace
TOP	Aboutness-shift topic operator/head
TOPP	topic phrase

TP	tense phrase
TR	transitive verb/suffix
TRL	translative case ('turn into')
u	addressee ('you')
[*u*D]	uninterpretable D-feature
[*u*φ]	uninterpretable phi- (agreement) features
VB	verbalizer, verbal suffix
VOICEP	voice phrase
VP	verb phrase
vP	"little v" phrase
WCO	weak cross-over
π, πP	person feature, person phrase
φ, φP	phi- (agreement) features, phi-features phrase
#, #P	number feature, number phrase

List of figures

Editors' note

Null subjects present a classic example of parametric variation, dividing languages into those that allow the subject of active finite clauses to remain silent, and those that do not. This book focuses on languages that are relatively understudied in this respect. It is a selection of twelve studies featuring null subjects in Slavic, Baltic, Finno-Ugric and Samoyedic language families. In particular, it includes chapters on East Slavic (Russian, Ukrainian, Belarusian), West Slavic (Polish, Czech), South Slavic (Bulgarian); on two Baltic languages, Lithuanian and Latvian; on Finnish, Hungarian, Mari and Saami from the Finno-Ugric language family; and finally, on two Samoyedic languages, Selkup and Nganasan, closely related to Ugric languages.

As the title of the book suggests, the main concern of this volume is to draw attention to null subjects in two less studied European language families, in particular, Slavic and Finno-Ugric. The two Baltic languages represented in the volume connect these two language families geographically, while the Samoyedic languages included here point towards Turkic and other Altaic languages, which also allow null subjects but are not represented in the volume.

The inventory is obviously incomplete. South Slavic languages are represented merely by Bulgarian, and Finnish and South Saami represent the Finnic branch. In this sense, the volume is not a concise handbook. Rather, it is a manifestation of the Null Subject Continuum. The collection is meant to be an appetizer, which will hopefully instigate further research.

As is usually the case with thematic volumes, the theoretical background of the chapters is not homogeneous. It ranges from minimalist, cartographic and conceptual semantic to historical, typological, corpus-based and descriptive approaches. These approaches, nonetheless, make it possible to compare the different methods used by these linguistic theories and to measure them against the data presented in the volume.

The volume starts with a Preface by **Anders Holmberg**, which, in addition to setting the scene, also serves as a recommendation of the volume to the linguistic audience.

In Chapter 1, **Jacek Witkoś** offers an overview of recent theories of null subjects within generative syntactic theorizing. Three new avenues of research have crystallized since Rizzi's (1982) seminal work on null subjects: (i) Holmberg's tripartite division of null subjects according to their projectional complexity (DP, φP, nP); (ii) Frascarelli's theory of licensing null subjects in the C-domain, as prescribed by the Topic Criterion; (iii) Barbosa's theory of the internal organization of null subjects as minimal nPs. These theories are often referred to in the chapters of the volume.

Egor Tsedryk in Chapter 2 discusses null subjects in East Slavic root clauses. He offers a comprehensive overview of the subject (null and overt) pronouns in Belarusian, Russian and Ukrainian, including (i) non-referential the third person plural generic pronoun, (ii) the second person generic pronoun and (iii) the reflexive marker used in generic statements. First of all, it is shown that referential null subjects are not linked to Topic operators in the left periphery. Pronominal elements are subsequently analyzed from the point of view of their morphosyntactic features and their categorial status. That is, null subjects are analyzed as projections of *phi*-features, in opposition to overt pronouns that have an additional D-layer in their extended nominal projection. The contribution of the D-head is twofold: (a) it signals the overt spell-out at the sensori-motoric interface, and (b) it ensures type shifting in the logical form. The chapter concludes with a list of spell-out rules operating in East Slavic languages.

In Chapter 3, **Nerea Madariaga** offers a unified account of referential null subjects in Russian both from the synchronic and the diachronic perspective. Building on previous work, she evaluates the crucial conditions for licensing null subjects in Russian according to the level of embedding (root vs embedded clauses), and verbal finiteness. Modern Russian, a partial Null Subject Language (NSL), displays a complex pattern: non-emphatic pronominal subjects in root clauses can or cannot be dropped, depending on (i) their informational interpretation (whether they are discourse Topics or successive occurrences of Topics in a topic chain), and (ii) additional syntactic restrictions involving locality requirements. In embedded clauses, null subjects are licensed under obligatory control in both finite and non-finite clauses, again, in absence of an intervener between the null subject and the left periphery. In Old Russian (a consistent NSL), on the contrary, every non-emphatic subject had to be dropped, in root and embedded, finite and non-finite clauses, regardless the distance of the antecedent in the text and despite the presence of potential interveners.

In Chapter 4, Polish null subjects are investigated by **Marta Ruda**. Polish is a consistent NSL with a basic SVO but in principle flexible word order, in which the interpretation of nominal arguments, including definite bare NPs, is, to a great extent, guided by contextual factors. It is shown that null subjects in Polish follow the same pattern, making available a much wider array of interpretations than what is observed, for example, in Italian, which is the paradigm case of consistent NSLs. With the theoretical focus on the left-peripheral approaches to the licensing and interpretation of null subjects, the author shows that the proposals based on the requirement that third person null subjects be identified with the (Aboutness-Shift) Topic heavily undergenerate when tested against the null subject facts of Polish. In general, the discussion leads to the conclusion that approaches which do not rely on information-structural notions for the licensing

of null subjects provide a more promising basis for furthering the theoretical understanding of the mechanisms underlying the phenomenon, including the patterns of cross-linguistic variation.

Chapter 5, by **Ludmila Veselovská**, investigates null subjects in Czech. The author first provides the Czech data to demonstrate the typical characteristics of a consistent NSL. Czech finite predicates, although located in a low position, carry rich agreement morphology allowing null subjects with a non-contrastive definite (specific) interpretation in all Tense-Aspect contexts for all three persons, both singular and plural. On the other hand, a generic (non-specific) subject in standard Czech requires either an overt proform *jeden* 'one' or a standard null subject. The chapter concentrates on Czech structures with (i) obligatory null subjects, and (ii) those which do not allow null subjects. The former includes agentless predicates, generic second person singular subjects, and anaphoric subjects in non-root clauses with a hierarchically symmetrical interpretation. The chapter proposes a complex agreement process including several independent checking domains for (i) the nominal *phi*-feature set, (ii) the D-feature set and (iii) a Topic/Focus feature. The theoretical framework used for the discussion and analyses is based on the diagnostics and claims made in the previous studies on null subjects.

In Chapter 6, **Dobrinka Genevska-Hanke** investigates Bulgarian, in the context of the existing classifications, as a consistent or a partial NSL. While providing ample empirical evidence for Bulgarian being a consistent NSL, the author shows that, in contrast to Italian (and thus contrary to expectation), Bulgarian allows overt subjects in Topic continuity contexts, implying that consistent NSLs are possibly subject to microvariation. Another reason for a more fine-grained classification is the availability of a generic null subject, characteristic of partial NSLs. While consistent NSLs are expected to use the reflexive *si*, Bulgarian (on a par with European Portuguese, a consistent NSL) allows for both of these options. There is evidence for a dissociation between the *si*-construction and genericity, on the one hand, and for the fact that *si*-constructions are formally similar (but not identical) to passive constructions, on the other hand.

Chapter 7, by **Axel Holvoet** and **Anna Daugavet**, deals with subjects without phonetic realization in Baltic languages, characterized in terms of agreement features on the verb, the type of syntactic environments they are compatible with, and the regular semantic interpretations associated with them. The discussion focuses on three types: (i) generic human null subjects that are masculine singular in terms of the verbal agreement, (ii) *they*-type null subjects that are masculine plural in terms of agreement (Siewierska's "episodic type"), and (iii) ambient (force) inanimate null subjects with causative verbs. Lithuanian and Latvian sen-

tences with null subjects show major differences in their use and in the different types of interaction with other (impersonal and passive) constructions. The areal connections of the constructions under discussion (Slavic and Finnic) are mentioned to clarify some points of interpretation.

In Chapter 8, **Urpo Nikanne** discusses licensing null subjects in Finnish active finite sentences. The focus is on argument structure and the analysis is based on the author's Conceptual Semantics framework. The licensing and interpretation of null subject sentences is explained as the interaction between three different levels of argument structure: (i) syntactic arguments (subject and object), (ii) lexically determined arguments ("logical" subject and object), and (iii) lexical conceptual structures. It is shown that empty syntactic subject is not needed for licensing a null subject if the system, as a whole, can link the structure to lexically determined arguments and a well-formed conceptual structure. Two kinds of NS are discussed: the referential (first or second person) null subject and the third person generic NS. In the plural, the latter is allowed only with verbs of communication. The author also provides an account of the different interpretive procedures applicable the respective null subject types.

Chapter 9 by **Gréte Dalmi** investigates the licensing conditions for the individual vs. generic reference interpretations of null arguments in Hungarian finite matrix and dependent clauses. The free occurrences of individual reference null arguments behave like R-expressions. The bound variable occurrences are subject to standard syntactic constraints such as locality, c-command and coreference. While lexical pronominal arguments with the individual interpretation require merely syntactic licensing, their null counterparts must be both syntactically and semantically licensed. Null arguments with the generic inclusive interpretation must be semantically licensed by the generic operator on the left periphery of the clause.

Generic reference null arguments also have free vs. bound occurrences in Hungarian. The free occurrences must always be lexical, just like *one* and *people* in English. The bound variable occurrences are always null and require a coreferential generic antecedent in some adjacent clause, but without the syntactic requirements applying to the bound variable occurrences of individual reference null arguments.

From the typological perspective, Hungarian resembles radical NSLs inasmuch that in finite clauses any argument of the predicate can remain silent as long as this argument can be reconstructed at some level of linguistic representation. Hungarian also differs from radical NSLs in disallowing free variation in the interpretation of the third person individual and generic reference null arguments. These facts place Hungarian somewhere between partial and radical NSLs in the typological scale of Null Subject Languages established by Holmberg & Roberts (2010).

Chapter 10, by **Jeremy Bradley** and **Johannes Hirvonen**, presents a corpus-based study of the conditions on licensing null subjects in Mari. Mari, a pluricentric language spoken in the Volga and Ural Regions of the Russian Federation, has two actively used literary norms – Meadow (Eastern) Mari and Hill (Western) Mari – and has historically been subject to Turkic and Russian language contact. First, the classic hypothesis that NSs are licensed by verbal *phi*-features is tested against an alternative hypothesis which attributes null subject licensing to a left-peripheral discourse-related head. Second, the authors perform a multi-level quantitative analysis to examine the importance of factors which do not have a categorical effect on subject expression, such as subject person (first, second, or third). As their empirical base, they use original fieldwork data, as well as the Corpus of Literary Mari recently compiled with the involvement of the authors. This corpus includes large bodies of texts from two literary standards of Mari dating back to before the Russian Revolution.

In Chapter 11, **Mikael Vinka** investigates two types of referential third person null subjects in South Saami, a Finno-Ugric language spoken in central Norway and Sweden by approximately 1,000 native speakers. The author addresses three distinct, but interrelated phenomena observed in South Saami, i.e. null subjects, logophoricity and cases of obligatory coreference in subjunctive complements. The general claim is that these phenomena have in common the fact that the occurrence of a null subject or a logophoric pronoun is contingent on the presence of a designated item in the C-domain which mediates the relation between the pronoun and its ultimate antecedent. In embedded contexts, two types of NS can be detected. One of them is *pro,* occuring in matrix clauses and indicative complement CPs. However, the null subject in subjunctive complements is incompatible with the *pro* analysis. On the basis of distributional facts, the author argues that this particular kind of null subject is a PF-deleted overt pronoun.

In Chapter 12, **Susann Fischer** and **Beáta Wagner-Nagy** investigate the distribution of referential and generic null subjects in Selkup and Nganasan, including the syntactic and pragmatic conditions determining their distribution. Nganasan got in contact with Russian relatively late, while Selkup has been under Russian influence for centuries. In Uralic languages (including Samoyedic languages), it is not obligatory to lexically express the subject and the object. Thus, these languages can be classified as *pro*-drop languages. At the same time, the circumstances that render subject omission possible have yet to be described extensively. Likewise, it is less known how the two closely related languages behave with regard to subject omission. In the course of the discussion, the authors seek to answer these questions.

Chapter 13, by **Piotr Cegłowski**, summarizes the main findings of the volume.

Acknowledgements

We wish to thank all the contributors for their patience and endurance in the course of compiling the manuscript. The technical editors and type-setters made a great effort to faithfully reproduce the data from the languages represented in the volume. We owe gratitude to De Gruyter's series editors and editors, who assisted us throughout the production of the volume.

We are indebted to the reviewers for their thorough and conscientious work, which helped the contributors immensely to make their chapters consistent, up-to-date, and to-the point. In particular, we wish to thank:

András Bárány (Postdoctoral Researcher, Leiden University, Leiden, Holland)
Joanna Błaszczak[†] (Professor, University of Wrocław, Wrocław, Poland)
Patricia Cabredo-Hofherr (Visiting Scholar, University of Surrey, United Kingdom)
Piotr Cegłowski (Professor, Adam Mickiewicz University, Poznań, Poland)
Federica Cognola (Associate Professor, Ca'Foscari University, Italy)
Gréte Dalmi (External Consultant, Hamburg University, Hamburg, Germany)
Roberta D'Alessandro (Professor, Leiden University, Leiden, Holland)
Mojmír Dočekal (Masaryk University, Brno, Czech Republic)
Anders Holmberg (University of Newcastle, Newcastle, UK)
Dara Jokilehto (Independent Researcher, Berlin, Germany)
Marit Julien (University of Lund, Lund, Sweden)
Olga Kagan (Assistant Professor, Ben Gurion University, Beersheba, Israel)
Iliyana Krapova (Professor, Ca' Foscari University, Venice, Italy)
Lidia Mazzitelli (Postdoctoral Researcher, University of Köln, Köln, Germany)
Lucie Medova (Postdoctoral Researcher, Masaryk University, Brno, Czech Republic)
Krzysztof Migdalski (Assistant Professor, University of Wrocław, Poland)
Edith Moravcsik (Professor Emerita, University of Wisconsin, Milwaukee, USA)
Urpo Nikanne (Professor, Åbo Akademi, Turku, Finland)
Genoveva Puskas (Professor, University of Geneva, Geneva, Switzerland)
Teodora Radeva-Bork (postdoc researcher, University of Potsdam, Potsdam, Germany)
Halldór Sigurðsson (Professor, University of Lund, Lund, Sweden)
Ida Toivonen (Professor, Carleton University, Ottawa, Canada)
Egor Tsedryk (Professor, Saint Mary University, Halifax, Canada)
Maria Vilkuna (Associate Professor, University of Helsinki, Helsinki, Finland)
Beáta Wagner-Nagy (Professor, University of Hamburg, Hamburg, Germany)
Jussi Ylikoski (Professor, University of Oulu, Finland)

Preface

The textbook and reference grammar view of null subjects as found in, for example, Finnish is that the pronouns are not needed because the agreement on the finite verb shows the person and number of the subject.

Because the personal endings sg1, sg2, pl1, pl2 attached to the verbs give clear information about the grammatical subject, it is natural that the corresponding pronoun subjects are often omitted (Karlsson 2018: 113).

In the case of languages with null subjects but no agreement, such as Chinese, the textbook view is that pronouns can be omitted because the context can indicate who the intended subject is, by providing a salient antecedent. This raises the question why that couldn't equally well be the case in languages with agreement. Indeed, there is important research from the late nineties onwards showing that contextual clues are relevant for the interpretation of null pronouns in languages with rich subject-verb agreement, too (Samek-Lodovici 1996; Grimshaw & Samek-Lodovici 1998; Frascarelli 2007; Cole 2010). It also raises the question why it couldn't be the case in languages like English and French, notorious for not employing null subjects. The answer is that it is the case. They make use of null subjects, too, as in *Can't tell you how happy I am to see you!*. As discussed by Haegeman (2013), this is not a marginal or exceptional phenomenon, but subject to strict universal, grammatical principles. Note how the 1st person singular pronoun is omitted in the embedded infinitival clause as well, but not in the embedded finite clause, a characteristic of this type of subject omission.

So it may be the case that all languages make use of null subjects. And why wouldn't they? It is a characteristic property of language use that meaning is conveyed by minimal means, observing computational efficiency, in Berwick & Chomsky's (2016: 101) terms. Their favourite example is the obligatory deletion (non-pronunciation) of all but one copy in a chain of copies derived by movement. Null pronouns is another case in point; if the meaning of an argument (chain) can be conveyed without any physical signal, then that is what happens, obligatorily or optionally, as the case may be. The conditions under which the meaning of an omitted argument can be conveyed is subject to cross-linguistic variation, though, which is what comparative research on null subjects/*pro*-drop wants to account for.

A fact which has become abundantly clear by virtue of much comparative research on *pro*-drop carried out ever since Perlmutter's (1971) seminal work is that there is a greater variety of null subject/*pro*-drop systems than was envisaged in the early days. The differences between *pro*-drop in languages with agreement and those without agreement have been an object of research from early on,

starting in earnest with Jim Huang's early work on Chinese (Huang 1984, 1989). The distinction between *pro*-drop of referential subject pronouns, as found in Italian, and expletive pronouns, as found in German and Icelandic (so called semi-*pro*-drop), was formally described in Rizzi (1982) (see Biberauer et al. 2010, ch. 4). There was some important research in the late 1980ies and 90ies on null subjects, particularly in Romance languages (Rizzi 1986; Barbosa 1995), but also in Greek (Alexiadou & Anagnostopoulou 1998) and in Arabic (Fassi Fehri 1993), among other languages, widely taken, at the time, to be archetypal agreement-based null subject languages. In Vainikka & Levy (1999), Modesto (2000a, b), Kato (2000), and a bit later, in work by myself and my colleagues, a distinction was identified between what came to be called consistent (or full) null subject languages and partial null subject languages (Holmberg 2005; Holmberg et al. 2009; Biberauer et al. 2010).

Consistent null subject/*pro*-drop languages are characterized by the following properties: (a) pronominal subjects are null unless they are focused/contrastive or introduce a new topic, (b) 3rd person pronominal subjects can be null in main clauses, interpreted by linking to an argument in the discourse context, (c) generic subject pronouns are not null, or are licensed by sentence-internal passive or reflexive morphology, and (d) they have a full set of person and number distinctions marked on the finite verb. In partial null subject/*pro*-drop languages, on the other hand, (a) pronominal subjects are only optionally null, (b) 3rd person pronominal subjects cannot be null in main clauses, (c) there is a null generic subject pronoun in active finite clauses, and (d) the subject-verb agreement paradigm is typically, though not necessarily, incomplete. The partial *pro*-drop languages that were identified and described in greatest detail were Brazilian Portuguese, Finnish, Hebrew, and Marathi. Several other Indo-Aryan languages probably belong to this class, as well (see Holmberg 2017 on Russian). In the references mentioned we also proposed a formal explanation of the dichotomy. The idea was that the nominal features of T, which end up as an agreement affix on the verb, include a definiteness feature in consistent *pro*-drop languages but not in the partial *pro*-drop languages. This has the consequence that a null 3rd person subject is necessarily interpreted as referential in consistent *pro*-drop languages, but not in partial *pro*-drop languages, where it can only be interpreted as generic, unless it has a referential antecedent, a controller, in a higher clause. First and second person null subjects are interpreted as such because they always have a contextual referent, the speaker and the addressee (to put it simply).

The third major type is discourse *pro*-drop languages (also called radical *pro*-drop languages). They have no nominal features in T. This has the consequence that a null subject can be interpreted as referential if it has an antecedent in the

immediate context, or, in the absence of an antecedent, as generic (Phimsawat 2011; Holmberg & Roberts 2013).

Subsequent research has questioned aspects of the theory, with the postulated trichotomy between consistent, partial, and discourse *pro*-drop. Barbosa (2019) argues that the partial *pro*-drop languages crucially share the property with discourse *pro*-drop languages that the null argument is not a pronoun but a minimally specified noun. The idea that discourse *pro*-drop is dependent on the absence of articles, and that argument drop in such languages is NP-ellipsis was proposed already by Jayaseelan (1999) and Tomioka (2003). The status of Russian in the trichotomy is controversial. Bizzarri (2015) has argued that Russian, counted among the partial null subject languages in Barbosa (2019) (on the basis of Lindseth 1998) and, more tentatively, in Biberauer et al. (2010), has some properties of a consistent Null Subject Language. Frascarelli (2018) and Frascarelli & Jiménez-Fernandes (2019) have shown by experimental research that there is more variation in the use of null subjects among speakers of partial (Finnish) and consistent Null Subject Languages (Italian, Spanish) than previously realized, and they propose a different formal account of the difference between these null subject types. Alexiadou & Carvalho (2018) discuss some striking differences between the two partial Null Subject Languages Finnish and Brazilian Portuguese. Sigurðsson (2011) articulates a theory where differences between *pro*-drop systems are explained as variation in how the null arguments are linked to the context. These are just some examples of recent work on null subject/*pro*-drop in the context of the consistent-partial-discourse trichotomy.

The reflections above concern null subjects in finite clauses. Null subjects are, of course, also found abundantly in non-finite clauses. In the 1980ies and 90ies this was regarded as governed by mechanisms distinct from those that governed null subjects in finite clauses. This is not the case anymore. Concomitant with the research on partial Null Subject Languages it has become clear that null subjects in embedded finite clauses are interpreted by rules similar to, though not necessarily identical to, control into non-finite clauses (Landau 2004; Biberauer et al. 2010: ch. 3; Sheehan 2018). This is not in contradiction with the consistent-partial-discourse trichotomy. Note, for example, how PRO, the null subject of infinitival clauses, is similar to the subject of finite causes in discourse *pro*-drop languages: referential if controlled by a referential antecedent, generic if not controlled, as in *It's nice [PRO to be rich]*.

Against this background, I find the present volume particularly welcome. What is the status of the various Slavic languages in the trichotomy of consistent, partial, and radical *pro*-drop languages? As mentioned, the position of Russian is controversial; in this volume it is scrutinized in two of the chapters. The West Slavic languages Polish and Czech show the outward signs of consistent Null

Subject Languages, as does South Slavic Bulgarian. Subjected to close investigations in three chapters in the book, they each turn out to have certain surprising properties, given the prevalent views of *pro*-drop sketched above. There is a chapter on impersonal generic null subjects in the Baltic languages Latvian and Lithuanian, two languages that I have always been keen to know more about. The Uralic languages, with the exception of Finnish and Hungarian, have not had their null subject-related properties examined in anything like the detail seen in the present volume, where, in addition to Finnish and Hungarian, there are chapters on Mari, South Saami, and two Samoyedic languages, Selkup and Nganasan. They all, in one way or other, present challenges for prevalent views on null subjects. Hungarian, for example, turns out to combine properties of three major types of NSLs. It is interesting, as well, to consider Urpo Nikanne's theory of Finnish null subjects, a representative of the school of thought where omitted arguments are radically absent from syntactic structure, being interpreted by semantic inference.

Anders Holmberg

References

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing AGR: word order, V-movement and EPP checking. *Natural Language and Linguistic Theory* 16 (3). 491–539.

Alexiadou, Artemis & Janayna Carvalho. 2018. The role of locatives in (partial) *pro*-drop languages. In Michelle Sheehan & Laura R. Bailey (eds.), *Order and Structure in Syntax II*, 41–67. Berlin: Language Science Press. https://langsci-press.org/catalog/book/115.

Barbosa, Pilar. 1995. *Null subjects*. Cambridge, MA: MIT. PhD dissertation

Barbosa, Pilar. 2019. *Pro* as a minimal NP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50 (3). 487–526.

Berwick, Robert & Noam Chomsky. 2016. *Why only us?* Cambridge, MA: MIT Press.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bizzarri, Camilla. 2015. Russian as a *pro*-drop language. Data and analysis from a new study. *Annali di Ca' Foscari* 49. 335–362.

Cole, Melvyn. 2010. Thematic null subjects and accessibility. *Studia Linguistica* 64 (3). 271–320.

Fassi Fehri, Abdelkader. 1993. *Issues in the Structure of Arabic Clauses and Words*. New York: Springer.

Frascarelli, Mara. 2007. Subjects, topics, and the interpretation of referential *pro*. *Natural Language and Linguistic Theory* 25 (4). 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null-subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.),

Null Subjects in Generative Grammar. A Synchronic and Diachronic Perspective, 211–239. Oxford: Oxford University Press.

Frascarelli, Mara & Ángel Jiménez-Fernandez. 2019. Understanding partiality in pro-drop languages: An information-structure approach. *Syntax* 22 (2–3). 162–198.

Grimshaw, Jane and Vieri Samek-Lodovici. 1998. Optimal subjects and subject universals. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*, 193–219. Cambridge, MA: MIT Press.

Haegeman, Liliane. 2013. The syntax of registers: Diary subject omission and the privilege of the root. *Lingua* 130. 88–110.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–64.

Holmberg, Anders. 2017. Linguistic typology. In Ian Roberts (ed.) *The Oxford Handbook of Universal Grammar*. Oxford: Oxford University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63 (1). 59–97.

Holmberg, Anders & Ian Roberts. 2013. The syntax-morphology relation. *Lingua* 2013 (130). 111–131.

Huang, C-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15 (3). 531–74.

Huang, C.-T. James. 1989. *Pro*-drop in Chinese: A generalized control theory. In Osvaldo Jaeggli & Ken Safir (eds.), *The Null Subject Parameter*, 185–214. Dordrecht: Kluwer.

Jayaseelan, Karattuparambil. 1999. Empty pronouns in Dravidian. In Karattuparambil Jayaseelan (ed.), *Parametric studies in Malayalam syntax*, 14–25. New Delhi: Allied Publishers.

Karlsson, Fred. 2018. *Finnish. A Comprehensive Grammar*. Abingdon and New York: Routledge.

Kato, Mary A. 2000. The partial *pro*-drop nature and the restricted VS order in Brazilian Portuguese. In Mary A. Kato & Esmeralda Negrão (eds.), *Brazilian Portuguese and the Null Subject Parameter*, 223–258. Frankfurt: Vervuert.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22. 811–877.

Lindseth, Martina. 1998. *Null-Subject Properties of Slavic Languages. With Special Reference to Russian, Czech and Sorbian*. München: Sagner.

Modesto, Marcello. 2000a. *On the identification of null arguments*. San Diego: University of Southern California, PhD dissertation.

Modesto, Marcello. 2000b. Null subjects without 'rich' agreement. In Mary A. Kato & Esmeralda Negrão (eds.), *Brazilian Portuguese and the Null Subject Parameter*, 147–74. Frankfurt: Vervuert.

Perlmutter, David. 1971. *Deep and Surface Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Phimsawat, On-Usa. 2011. *The Syntax of pro-drop in Thai*. Newcastle: Newcastle University. PhD dissertation.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–557.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. New Brunswick, NJ: Rutgers University. PhD dissertation.

Sheehan, Michelle 2018. On the difference between exhaustive and non-exhaustive control. In Federica Cognola & Jan Casalicchio, eds.), *Null Subjects in Generative Grammar. Synchronic and Diachronic Perspectives*, 141–170. Oxford: Oxford University Press.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42 (2). 267–304.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In Kerstin Schwabe & Suzanne Winkler (eds.), *The Interfaces Deriving and Interpreting Omitted Structures*, 321–339. Amsterdam: John Benjamins.

Vainikka, Anne & Yonota Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17. 613–671.

Jacek Witkoś

1 Pro subject(ed) to challenge: The concept of the null subject and typologies of Null Subject Languages

1 Introduction: *pro* enters the stage

In the Principles & Parameters framework, null subjects in finite clauses were particularly interesting from a cross-linguistic perspective, as they seemed to be attributable to a single parameter separating Null Subject Languages (NSLs), such as Italian and Spanish, from non-NSLs, which require lexical or pronominal subjects (e.g., English and French). This view was in part driven by early attempts to identify a set of properties that characterize NSLs (Taraldsen 1980; Rizzi 1982) and the subsequent attempt to offer a generalized theory of null pronouns, so-called "little *pro*'s", in both subject and object positions (Rizzi 1986). The theory was subsequently extended to "impersonal sentences" and "non-referential" instances of null arguments, the so-called "arbitrary *pro*" (Cinque 1988; Cabredo-Hofherr 2003). Furthermore, it has been shown that in addition to purely formal means to license null subjects (i.e. verbal agreement), grammar may also resort to discursive means (e.g., topicalization) that play a crucial role in "radical *pro*-drop" languages such as Japanese and Chinese (Huang 1984, 1989). These early attempts triggered an unprecedented interest in null arguments across languages. Subsequent research has shown that null subject phenomena are not reducible to a single parameter.

The original Null Subject Parameter was formulated in Rizzi (1982: 142) and involved a set of two parametric options:

(1)   *The Null Subject Parameter*
    a.   INFL can be specified [+pronoun].
    b.   INFL which is [+pronoun] can be referential.

The parametric option in (1a) distinguishes between non-NSLs (e.g., English, French, Swedish), so languages that force lexicalisation of pronouns in the subject position under all syntactic and morphosyntactic circumstances: referential, impersonal, expletive, etc., and languages that admit the option of dropping the subject pronoun (e.g., Greek, Spanish, Italian, Polish, Russian, Finnish, Hebrew, Icelandic, Chinese, Japanese, Korean, etc.). Parameter (1b) distinguishes between NSLs that allow for the drop of only non-referential (expletive) or qua-

si-referential (subject of weather predicates) subject pronouns and the ones that license null subjects referring to both referential and non-referential nominals.

As an empty category, *pro* fits well with the then classical paradigm of (nominal) empty categories presented in Chomsky (1982):

(2) *The GB paradigm of empty categories*

	pronominal	anaphoric
PRO	+	+
pro	+	−
NP-trace	−	+
variable	−	−

With empirical studies of different NSLs accumulating, the definition of the parameter changed accordingly:

(3) *Null Subject Parameter* (Rizzi 1986: 518–523)
   a. *pro* must be licensed
       Licensing: *pro* is case-marked by X^o_y, where *y* is parametrized.
   b. *pro* must be identified
       Identification: *pro* inherits the φ-feature values of X^o_y (if it has φ-features; if not, *pro* gets a default interpretation, typically *arb*).

On the strength of (3), null pronouns could come in two types (Rizzi 1986: 543):

(4) a. An NP is referential only if it has the specification of person and number.
     b. An NP is argumental only if it has the specification of number.

These two types are not independent from each other but they form a hierarchy and a set/subset relationship: the set of all argumental *pro*s is specified for number and a subset of this set is additionally specified for person. Rizzi's identification is equivalent to today's feature valuation: T values the φ-features of *pro*. In the structure in (5), *pro* is formally licensed by T, assuming that T belongs to the defined and parametrised class of designated licensers (but other heads, e.g. v^o, can also license null arguments):

(5) $[_{TP}$ pro$_i$ $[_{T'}$ T$_{i[3PL]}$ . . .]]

The emerging picture was that rich inflection of the verb largely facilitated the drop of the subject pronoun and correlations between these two phenomena were soon established. Yet, the interdependence between the morphological composition of

the inflected verb and appearance of *pro* was not as straightforward as one would have wished. For instance, certain languages with a rich inflectional system licensed *pro*-drop, while others with relatively rich inflection did so only in a very limited set of cases. Still others, which showed no person/number distinction, did allow for the null subject (and the null object, for that matter). Consider the following verbal paradigms in Spanish, German and Japanese (Jaeggli & Safir 1989: 27–29):

(6)

Spanish			German		
habl-o	I speak	1SG	*(ich) arbei-te*	I work	1SG
habl-as	you speak	2SG	*(du) arbeit-est*	you work	2SG
habl-a	he/she speaks	3SG	*(er/sie) arbeit-et*	he/she works	3SG
habl-amos	we speak	1PL	*(wir) arbeit-en*	we work	1PL
habl-áis	you speak	2PL	*(ihr) arbeit-et*	you work	2PL
habl-an	they speak	3PL	*(sie) arbeit-en*	they work	3PL

(7)

Japanese	
yom-ru	read-present
yom-ta	read-past
yom-anai	read-neg
yom-eba	read-conditional
yom-oo	let's read
yom-itai	want to read
yom-are	was read
yom-ase	make read

While Spanish allows for thematic and expletive null subjects, German allows only for the expletive ones. Yet, the fact that the inflectional paradigm of the verb includes one or more syncretic forms cannot be the only reason, as Irish, for instance, also allows for syncretic and thematic null subjects. Jaeggli & Safir proposed that it is not plain richness of the paradigm and number of syncretic forms that matters but the uniformity of the paradigm:

(8) *The Null Subject Parameter* (Jaeggli & Safir 1989: 29)
Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.

The paradigm counts as uniform if it contains only morphologically complex forms or only morphologically simple forms; "mixed" paradigms do not count as uniform and they do not license the null subject.

(9) *Morphological Uniformity* (Jaeggli & Safir 1989: 30)
An inflectional paradigm P in a language L is morphologically uniform iff P has either only underived inflectional forms or only derived inflectional forms.

By introducing the Null Subject Parameter and the notion of the uniformity of the morphological paradigm, they set the scene for the scholarly discussion of the null subject phenomenon for two decades. In order to accommodate various types of null subjects they kept the notions of the licensing of *pro* and its identification distinct. Paradigm uniformity served to license *pro*, while its identification (as a thematic or non-thematic *pro*) depended on one of the three strategies. The first one involves agreement with the T/Agr node, another one a c-commanding T/Agr or a c-commanding DP from the superordinate clause, and the third one a variable bound by a null operator, rather than *pro*.

Identification by agreement is regulated by the following principle:

(10) *Identification by Agreement*
AGR-Tense can identify an empty category as *pro* iff AGR-TENSE (case) governs the empty category.

One of the consequences of this definition is that the features of T/Agr must act jointly as one head. According to Jaeggli & Safir, this is the case of the morphologically uniform Romance *pro*-drop languages, where these features are placed in Infl, in line with Chomsky's (1986) *Barriers* system. Specifically, they submit that the V-2 Germanic languages cannot license argumental *pro* once it is assumed that the property of T is placed in C, while AGR is placed in Infl. Thus, a thematic null subject cannot be licensed in German, despite its relatively rich inflection in (6), as there is no single head that case-governs the subject position and mediates in subject-verb agreement. They point out that West Flemish, where the complementizer is inflected (so both AGR and T are in C), does allow for thematic null subjects.[1]

As for Chinese and Japanese, which show no distinct AGR/T marking on the verb, they assume that the null subject is licensed through a chain of local c-command relations, as proposed in Borer (1986, 1989): the subject in the main clause

---

1 The inflected complementizer in West Flemish licenses a null subject in the embedded clause, while the uninflected one does not:

(i) *dasse*   pro   *komt*
     that   (she) comes
(ii) **da*   pro   *komt*
     'that (he/she) comes...'

passes its φ-features onto the AGR in its clause, and these features get transmitted to the embedded INFL/AGR under c-command. Subsequently, *pro* in the embedded clause is identified through its (silent) INFL/AGR with the features of a c-commanding NP (typically the subject). This classic view of the Null Subject Parameter was successfully developed and adapted in many accounts of particular languages until the onset of the new minimalist methodology.

2 The spectrum of the null subject phenomena

The current scholarly depiction of the *pro*-drop phenomenon is dramatically distinct from the one commonly adopted in the 80s. A number of subtypes of NSLs have been identified and a number of divisions and subdivisions have been proposed, many of these are addressed by the contributions in this volume. Barbosa (2019: 487–488) uses the following classification of NSLs:

(11) Classification of NSLs:
   a. Languages with rich subject agreement morphology (henceforth consistent NSLs, after Holmberg 2005), such as Italian and Greek; subjects are freely dropped under the appropriate discourse conditions.
   b. Languages with agreement and referential null subjects whose distribution is restricted (partial NSLs), such as Hebrew, Finnish, Russian and Brazilian Portuguese.
   c. Languages that lack agreement, such as Chinese, Japanese and Korean. These have been described as allowing any argument to be dropped, not just subjects. They will be labelled here discourse *pro*-drop languages.
   d. Languages that have only impersonal and quasi-argumental null subjects (Icelandic, Faroese and a range of Creoles), generally referred to as expletive *pro*-drop or semi-*pro*-drop languages.

One large-scale research question is how the distinctions above can be captured by the Null Subject Parameter in (1), (8) or any of its current equivalents. In reaction to the criticism of the classical GB notion of parameter setting expressed for example in Newmeyer (2004), Biberauer (2010, 2018) attempts to consistently classify NSLs using the notion of a parameter hierarchy, which is a set of related interlinked parametric choices following the pattern:

(12) NONE   > ALL   > SOME

The pattern reflects the following set of options:

(13)   Does P(property) characterise L(anguage)?
    a.   Macroparameter: NO or Yes (Do all relevant heads have P?)
    b.   Macroparameter: YES or NO (Does a natural subset of heads have P?)
    c.   Mesoparameter: YES or NO (Does a restricted natural-class subset of heads have P?)
    d.   Microparameter: YES or NO (Do only lexically specified items have P [nanoparameter]?)

Biberauer observes that the concept of feature hierarchy and distinctions within it shown already in (13) leads to a specific classification of types of INFL, making two fairly innocent assumptions, i.e. (a) any INFL licensing a null subject must be pronominal, and (b) the features present on *pro* must match those on INFL, Biberauer (2018: 112):

(14)   a.   [pronominal], [person], [number] → referential *pro* [pronominal], [person], [number]
    b.   INFL [pronominal], [person] → quasi argumental pro [pronominal], [person]
    c.   INFL [pronominal] → expletive pro [pronominal]
    d.   INFL → no null pronouns

The four types of INFL distinguished on the basis of its feature composition lead to the definition of three types of NSLs (plus a non-NSL type) in her work:

(15)   a.   Full (consistent) NSL: licenses both referential and non-referential (i.e. expletive) null pronominals, e.g. Italian, Spanish, Greek.
    b.   Expletive or semi-NSL Type I: only licenses null non-referential pronominals, i.e. so-called expletives, which include both the quasi-argumental ('weather') type, and the non-argumental, genuine 'placeholder' type, e.g. Icelandic, Yiddish
    c.   Expletive or semi-NSL Type II: only licenses null non-argumental pronominals, i.e. only the non-argumental expletives, e.g. Dutch, German.
    d.   Non-NSL: does not license null pronominals (*pro*) at all, e.g. English, French.

The content of the NSL groups is not random and accidental but can be credited
to the following parameter hierarchy, respecting the guidelines in (14) above:

(16)   Is INFL pronominal
        (rich AGR)

NO   YES
(English, French)   (Does distinctively marked [number] show on INFL?

              NO   YES
        (German, Dutch)   Does distinctively marked [person] show on INFL

                        NO   YES
                  (Icelandic, Yiddish)   (Italian, etc.)
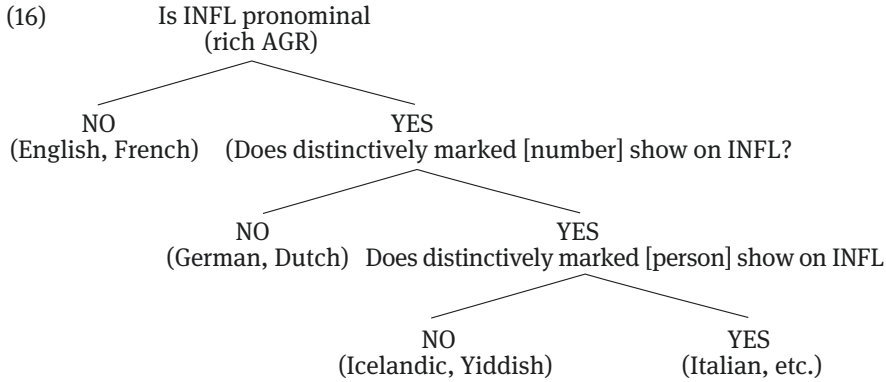
As the current volume concentrates on the discussion of the null subject phenom-
ena in Slavic and Finno-Ugric sub-families, we shift our attention to key general
properties of *pro* showing up there. As it happens, these languages are typically
either consistent or partial NSLs. Out of the four groups in the classification in
(11), the consistent and partial NSLs appear to be the closest cousins; both allow
for NSLs in main and embedded clauses. Yet, upon closer scrutiny several proper-
ties distinguish one group from the other, such as the interpretation of the lexical
pronominal subject in the subjacent embedded clause and the effect of third-
party subject intervention.

(17) *O João*ᵢ *disse* *que* (*ele*ᵢ) *tinha*   *comprado* *uma* *casa*.
      the João   said   that   he   have.PST.3SG bought   a   house
      'João said that he had bought a house.'
      (Brazilian Portuguese, Holmberg & Sheehan 2010: 131)

The example above shows licensing of the null subject in the subjacent embedded
clause in Brazilian Portuguese. In the European variant of this language (Euro-
pean Portuguese), the lexical subject pronoun *ele* 'he' is shunned in this position,
unless it is used contrastively. European Portuguese is a consistent NSL and the
lexical pronoun in the subject position is used for a reason, the null form being
the expected "run-of-the mill" standard. In the variety of Portuguese spoken in
Brazil, there appears to be free variation in this respect, both the null subject
and the lexical pronoun can be used interchangeably, without any difference in
interpretation. A similar free variation in this context between the null pronoun
and its lexical equivalent is reported for Finnish, Marathi, Russian and Hebrew.

An intervening third party subject also seems to affect the distribution of null subjects differently in the varieties of Portuguese, and in consistent and partial NSLs, in general:

(18) *O João*ᵢ *disse* [*que os moleques*ⱼ *acham* [*que* *(ele*ᵢ) *é* *esperto*]].
the João said that the boys believe that he is smart
'João says that the children believe that he is smart.'
(Brazilian Portuguese, Holmberg & Sheehan 2010: 137)

In consistent NSLs the presence of a distinct intervening subject does not adversely affect the subject pronoun drop, so example (18) would be grammatical in European Portuguese. In partial NSLs, such intervention blocks the null subject option and a lexical pronoun (*ele*) must be used in Brazilian Portuguese.

Partial NSLs may also display differences in the licensing of the null subject in the main clause with respect to grammatical person: for example, Finnish and Hebrew allow for the lexical drop of the referential 1ˢᵗ and 2ⁿᵈ person pronouns but not the 3ʳᵈ person pronouns.² In all partial NSLs the 3ʳᵈ person null subject in main clauses can be interpreted inclusively, as an impersonal pronoun, that is embracing in its denotation both the speaker and the addressee. Somewhat surprisingly, this option is not available in consistent NSLs, where the inclusive impersonal interpretation of the verb marked for 3ʳᵈ person requires lexical support in the form of the reflexive clitic:³

(19) *È* *assim* *que* *faz* *o doce*.
is.ꜱɢ so that make.3ꜱɢ the cake
'This is how one makes the cake.'
(Brazilian Portuguese, Rodrigues 2004: 72, cited in Barbosa 2019: 490)

(20) *È* *assim* *que* *se* *faz* *o doce*.
is.3ꜱɢ so that SE make.3ꜱɢ the cake
'This is how one makes the cake.'
(European Portuguese, Barbosa 2019: 490)

---

2 The extraordinary status of the 1ˢᵗ and 2ⁿᵈ person null subjects is discussed further in Section 1.3.
3 Consistent NSLs allow for null subjects with exclusive impersonal interpretation; for instance, Polish uses 3PL form for this purpose:

(i) *Mówią* *że* *benzyna* *stanieje*.
say.3PL that petrol become.cheaper
'They (people) say that petrol will become cheaper.'

Further differences between the two NSL groups become evident and are extensively studied (some aspects thereof are mentioned below). Holmberg & Sheehan (2010: 132) provide a comparison of basic features of consistent NSLs, partial NSLs and non-NSLs:

(21)   The null subject typology

	Consistent NSL	Partial NSL	Non-NSL
Example languages	Spanish, Greek, Turkish, Arabic, Telugu	Brazilian Portuguese, Finnish, Marathi	English, French, Sorbian, Somali
Non-root, non-controlled 'topic' pronouns	null	overt	overt
Inanimate subject pronouns	often null	overt/null if controlled	overt
Controlled subject	null	null/overt (optional)	overt
Generic subjects	overt	null	overt
Non-thematic subjects	null	null	overt

Every subsequent chapter in this monograph contains references to at least one of the properties enumerated above.

3 What is (left of) *pro* in 21st century linguistic research?

This section contains a subjective review of major current views on the nature, licensing and interpretation of the null subject and centres around three major approaches: the Holmberg/Roberts approach, formulated in a series of publications, underscoring the role of the morphosyntactic factors in the derivations with null subjects, the approach based on the key role of topic elements placed in the articulated left periphery of the clause (Frascarelli 2007, 2018), where the (silent) topic provides the null subject with its reference, and the approach where *pro* corresponds to a minimal nominal element nP/φP similar to null objects and licensed through semantic predicate modification (Barbosa 2019). The three major approaches above have been formulated almost concurrently and they largely feed on one another, so they are intertwined and interconnected to a large degree. The impression one formulates when properly introduced to them, is that they all recognise the common core of the null subject problem (the licensing of the null

subject in consistent NSLs) but they differ either in the execution of the central idea or in explaining its variations applied to more peripheral cases. All in all, this synergy in research efforts has led to the emergence of a new paradigm in the studies on NSLs. It appears that the sanctified relics of the GB era *pro* have been excavated and subjected to detailed re-examination in line with the minimalist methodology. The results are fascinating, as it turns out that further studies of the null subject and its properties bring about new arguments shifting the scales in favour of one of the three accounts or another.

3.1 Holmberg/Roberts' approach

In Holmberg/Roberts' approach, *pro* falls victim to incorporation or copy deletion. Foundations for current minimalist analyses of the null subject were laid in Holmberg (2005). Homberg observes that the minimalist procedure of feature valuation based on Agree, where T bears unvalued uninterpretable φ-features and acts as probe accessing a DP goal bearing interpretable and valued φ-features, as defined in Chomsky (2000, 2001), is incompatible with the GB-era notion of licensing of null subjects in (1). Chomsky's view on the feature composition of finite T practically turns tables on Rizzi's (1982, 1986) analyses in (1) and (3) above; instead of finite T identifying and licensing *pro* (providing its content), it is the DP goal (*pro*) that is supposed to value the features of T. Holmberg proposes that distribution of null subjects to a large degree hinges on the properties of T in the narrow syntactic derivation. Rizzi's notion of *pro* cannot be upheld any more, and is replaced by two *pro*-related constructs, i.e. either a set of nominal φ-features incorporated into T or a pronoun that undergoes deletion in Spec, TP of an embedded clause because it is controlled/bound by a DP argument placed in the superordinate clause. At least the first embodiment of the traditional *pro* is clearly compatible with Rizzi's (1a): T becomes pronominal if nominal features incorporate into it. The other variety of *pro* looks somewhat similar to its well-known cousin, an obligatorily controlled subject of infinitives and gerunds known as PRO.

Holmberg (2010a: 95) schematically presents the relationship between T and the pronominal subject in the following manner. The head T bears a D-feature, which is taken to be interpretable and valued in Holmberg (2005), but unvalued in Holmberg (2010), a difference with considerable consequences for the derivation, as shown below:

(22) $[_{TP} [_{T'} T_{[u\varphi, D]} [_{vP} \varphi P_{[3SG]} [v' \ldots .]]]]$

Holmberg's major point of interest is defining the distinction between consistent and partial NSLs, so a distinction somewhat related to Rizzi's parameter (1b). He introduces a key distinction between the two language types through feature composition of finite T. In consistent NSLs, T, as expected, bears a set of unvalued φ-features but also a D-feature, while in partial NSLs, T does not bear the D-feature.

Null pronouns in consistent and partial NSLs come in several types. Let us first examine the former subset of NSLs.

In consistent NSLs, null subject pronouns have impoverished structure and correspond to weak pronouns in the classification of Cardinaletti & Starke (1999). They are specified for φ-features and an unvalued D-feature (according to Holmberg 2005), so they cannot independently refer to a pronoun or a group. Holmberg calls them φPs, a term borrowed from Déchaine & Wiltschko (2002). They value the φ-features of T, while T values the feature of D, because its D-feature is interpretable.[4] The lexical pronouns, which are DPs, bear a valued feature D and a referential index (just like phrasal DPs). When used in a particular derivation, they value the φ-features of T, as well as its D feature, in a predicable manner, raising to Spec, TP on account of the EPP.

The mechanics of the valuation of the D feature on T undergoes a significant modification in Holmberg (2010a,b) with respect to consistent NSLs. Here *pro* is still a weak pronoun, a minimal φP containing valued φ-features involved in a Agree relation with T bearing a corresponding set of unvalued φ-features but also an unvalued D feature (valued by the aboutness shift topic placed within the CP area of the clause, see below). The valuation of the φ-features of T consists in copying the values from the goal φP and setting them on T. As these features are the only content of φP, it appears that the copying procedure in fact copies the entire φP and merges it with T. Holmberg follows Roberts (2010) and takes such exhaustive copying of the features onto the probe T to result in chain formation, with T (or rather its φ-feature content) forming the head of the chain and its tail, i.e. φP in Spec,vP. As the head of the chain is typically pronounced, T is pronounced (as a suffix on V). Thus, *pro* in a consistent NSL is a weak pronoun practically incorporated into T.

---

4 To be precise, Holmberg (2005: 556) defines the null subject of consistent NSLs in the following manner: "We may assume that the relation between φP and D is an Agree relation: φP has a feature [*u*D] (unvalued D) that is valued either by merging D with φP, which yields [$_{DP}$ D φP], or, by merging D as a component of I in a local c-commanding relation to φP merged with vP (where φP usually ends up remerged with IP). If φP with its [*u*D] features is not locally c-commanded by D, it can still be licit if it is bound by a DP (subject to parametric variation), or interpreted as generic. Valuing [*u*D] precludes the generic interpretation of the null subject. Therefore, consistent NSLs have to resort to a variety of 'overt strategies' to express the meaning of a generic subject pronoun."

Following Samek-Lodovici (1996) and Frascarelli (2007), Holmberg intro-duces another element into the licensing of the D-feature and *pro*: a particu-lar type of topic called the aboutness-shift topic. The following example from Samek-Lodovici (1996) illustrates the impact of the topic on the interpretation of a null subject in a consistent NSL language:

(23) a. *Questa mattina,   la mostra   è stata visitata   di Gianni.*
        this morning   the exhibition was   visited   by Gianni
        Piu tardi   *__/egli/lui   ha   visitato   l'università.
        later   (he)/he/he   has   visited   the university.
        'This morning the exhibition was visited by Gianni. Later he visited
        the university.'
     b. *Questa mattina. Gianni   ha   visitato   la mostra.*
        this morning   Gianni   has   visited   the exhibition.
        Piu tardi   __ ha   visitato   l'università.
        Later   has   visited   the university
        'This morning Gianni visited the exhibition. Later he visited the university.'
        (Samek-Lodovici 1996, cited in Holmberg 2010a: 96)

Null subject is impossible in (23a), although the NP antecedent to it in the *by*-phrase is relatively prominent. A lexical pronoun must be used, as it must introduce a new aboutness-shift topic, with the subject of (23a) 'the exhibition' functioning as the aboutness-topic up to this point. In (23b) the subject of the main clause serves as the aboutness-topic, so a referential *pro* in the embedded clause can be licensed. Thus, the menu of choices for the lexical antecedent of *pro* in the consistent NSL is then said to be determined by the topic. Acknowledging Frascarelli (2007), Holm-berg (2010a, b) assumes that the (silent) Aboutness Topic (A-Topic) is always syn-tactically represented in the C-domain of consistent NSLs and it serves as anteced-ent for the null subject. The A-Topics may form a chain and the referential index of this topic chain is identical with the index/value of a lexical DP present in preced-ing discourse. Schematically, the licensing of a referential *pro* looks as follows:

(24) $[_{CP}$ <*Gianni$_1$*> [*questa mattina Gianni$_1$ ha visitato la mostra*]]
      $[_{CP}$ <Ø$_2$> [*piu tardi ha* φP$_2$ *visitato la mostra*]]
      1 = 2
      (Holmberg 2010a: 96)

He takes the antecedent of the null subject in consistent NSLs to be a (possibly null) Aboutness-Shift Topic, base generated in the CP area (articulated left periphery). The Aboutness-Shift Topic ultimately refers back to a lexical DP spelled out in the preced-

ing discourse. Technically, the topic also bears a valued D feature, which it uses to value the D-feature of T. This valuation consists in copying the referential index of the topic onto T. At the same time, the EPP property of T is also checked by the topic, which implies that the φP pronoun, i.e. *pro*, could in principle remain in its base position of Spec,vP, where T reaches it under Agree to have its φ-features valued.[5]

The reason for which consistent NSLs do not have an inclusive impersonal *pro* (see, e.g., (21)) is that this type of defective pronoun (φP) could never value the D feature on T on its own and the obligatory presence of the aboutness topic excludes the impersonal (indefinite) interpretation.

The derivational circumstances are somewhat different in partial NSLs. Holmberg assumes that here the aboutness-shift topic is not licensed and T in these languages does not bear the D feature. Yet, it still has the EPP property. These pronominal null subjects fall into two types.

Thus, Type I weak pronouns bear only φ-features and no D-feature. According to Holmberg (2010a: 105), they value the φ-features, in doing so form a chain and incorporate into T, but they can only have generic interpretation, since neither this weak pronoun nor T bears the D-feature:[6]

(25) . . . [$_{CP}$ *että* [$_{TP}$ *tässä* [$_{T'}$ *istuu*+T$_{3SG,EPP}$ [$_{vP}$ φP$_{3.SG}$ *<istuu> mukavasti <tässä>*]]]]
        that   here   sits   comfortably
   '. . . that one can sit comfortably here.'

Type II weak pronouns are said to bear not only a set of valued φ-features but also an unvalued D feature ([*u*D]), which allows them not only to value the φ-features of T but also to move up to Spec,TP:

---

5 Holmberg (2010a: 95) admits that his earlier idea that finite T in consistent NSLs should bear a valued and interpretable D feature had to be abandoned in favour of the T bearing [*u*D] for two reasons. One was the dependency between the null subject and the A-Topic, which was not considered in Holmberg (2005), and the other was that some of the consistent NSLs have indefinite overt subjects that participate in Agree with T. But this should lead to a conflict of the D feature values. T bearing [*u*D] solves these problems and results in appropriate interpretations, coupled with the workings of the A-Topic.

6 Holmberg (2005: 552) defines the null subject of partial NSLs in the following manner, where the presence or absence of the [*u*D] feature does not apply: "I propose that the bound null pronoun and the generic null pronoun are the same category. I propose, furthermore, that it is φP. Since it values Agr [. . .] it must have inherently valued (hence interpretable) φ-features. Crucially, though, it lacks the substructure required for a definite, referential category. In the spirit of Longobardi (1994), I take the lacking property to be the head D, in the absence of which the pronoun cannot refer to an individual or group, either independently/deictically or under coreference with an independently referring DP. However, it can be a variable bound by a QP or a DP in a higher clause. As a last resort, it can be interpreted as generic."

(26) *Jari*₁ *sanoo* [CP *että* [TP uDP [T' *istuu*+T₃.SG,EPP [vP φP₃.SG <*istuu*> *mukavasti tässä*]]]]
Jari says that sits comfortably here
'Jari says that he sits comfortably here.'

This step is crucial, for it makes them accessible to an argument in the higher clause which controls them; this specific relation of control is all-important, for it provides for the definite interpretation and index/value of the 3rd person null pronoun. This relation, though technically distinct from obligatory control of PRO (Hornstein 1999, 2001; Landau 2000, 2006), is sensitive to intervention, see (17) above, and relatively local. No null topic mediates in this relation, so the lexical controller directly accesses the null pronoun. Still, for Holmberg the weak pronoun φP (a successor of *pro*) derivationally functions like any other syntactic object of the minimalist vintage; it is first-merged in a thematic position Spec,vP, it bears features that need to be valued (accessed and copied in a minimal domain if the probe) and it is subject to external and internal merge. It is pronounced as the suffix on the verb in consistent NSLs.

Ultimately, the only way to express a definite pronoun in the main clause of a partial NSL is to run a derivation including a pronoun with its own valued D-feature. It is externally merged in Spec,vP, values the φ-features of T and raises to Spec,TP to satisfy the EPP:

(27) a. *Hän on ostanut uuden auton.*
he has bought new car
'He has bought a new car.'
b. [TP *hän*[D1,3SG,NOM] [T' *on*+T[3SG, EPP] [vP < *hän*[D1,3SG,NOM]> *ostanut*]]]

As for non-NSLs (e.g., English and Swedish) the subject is never incorporated, so probably the pronouns in these languages can only function as DPs, never the weak variant of the φP. But Holmberg (2010a) recognises the need for preventing the option open to type two null subjects in partial NSLs, which may be DPs with [uD]. Within the approach he has devised, this option would mean that English and Swedish could control/bind null subjects in tensed embedded clauses, which they obviously do not. He proposes the following postulate that solves this problem:[7]

---

7 As pointed out in Holmberg (2005, 2010a), a principle like (28) is more of a postulate, as it must allow for a principled set of exceptions, such as 1st person diary drop or conjunction reduction.

(28)   T has a [P]-feature, where P is a feature requiring lexicalisation of the category merged in [spec, TP]

Now, for the broader picture. When painting a more comprehensive portrait of null subject phenomena across the linguistic landscape, Holmberg (2010a,b) Roberts (2010), Holmberg and Sheehan (2010) and Roberts & Holmberg (2010) focus on three properties involving parametrised options. First, does a language have a P-feature on T or not? This choice determines the [+NSL] or [-NSL] status of a given language, Next, does a language have a [*u*D] feature on T or not? Selection of either option produces either a grammar of a consistent NSL or a partial NSL. The third option concerns the mode of the satisfaction of the EPP property of T, i.e. whether or not it is φ-related. In other words, can only nominal elements (including null φPs and the A-Topic) be merged with T to form its specifier or will any other category do as well? Holmberg (2010a: 121–24) ends up reviewing a few sets of hierarchies of parametric options concerning NSLs, of which the hierarchy below seems to be a representative example:

(29)   Hierarchy of null subject parameters:
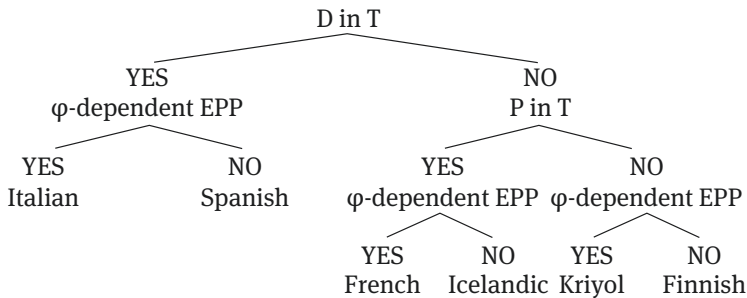
Holmberg & Sheehan (2010) present a detailed account of the interpretation of null subjects in embedded clauses in consistent and partial NSLs. They insist that in the latter type of NSLs the null subject is controlled by a c-commanding argument placed in the main clause. They forcefully argue that the relation between the DP antecedent of *pro* in the main clause and *pro* in the embedded clause shares basic similarities with obligatory control of PRO (e.g., Landau 2000, 2006; Hornstein 1999, 2001), namely locality restrictions (the next clause up) and c-command.

Yet, apart from the similarities there are also differences between these two relations. One concerns the notion of the independent tense. Landau (2000, 2006) observes that the temporal specification of the controlled infinitive can be either simultaneous with the event denoted by the verb of control (the case of

exhaustive control) or subsequent to the event denoted by the control predicate (the case of partial control). Yet, this limitation does not seem to apply to controlled *pro*, as in the following example from Marathi, the event in the embedded clause is independent and previous to the main clause:

(30) *Seema₁ kabul karte aaj ki (ti-chya-ni₁) kal chuk*
Seema agree does today that she-ERG yesterday mistake
dzali.
make
'Seema admits today that she made a mistake yesterday.'
(Marathi, Holmberg & Sheehan 2010: 136)

Another difference pointed out by Holmberg & Sheehan is breach of locality in this control relation; the controller DP is allowed not to appear in the next "clause up" in some partial control NSLs if the intervening clause does not contain any argument at all:

(31) *A Maria₁ disse* [*que é verdade* [*que (ela₁) entornou o copo*]].
The Maria said that is true that she knocked over the glass
'Maria said it's true that she knocked over the glass.'
(Brazilian Portuguese, Holmberg & Sheehan 2010: 137)

Control of *pro* also differs from obligatory control of PRO in requiring unambiguous c-command by the controller and it excludes a possibility open to PRO of being controlled through connectivity or "logophoric extension" (Landau 2000), where the possessor of certain NPs, such as *plan*, *career*, etc., is allowed to control PRO, though c-command does not hold:

(32) *It helped* [*John's career*] [PRO *to have an uncle on the board*].

In the example above *John* does not c-command PRO, yet obligatory control of PRO is allowed. Surprisingly, control of *pro* is typically disallowed under such circumstances:

(33) *O plano do José₁ realiziou-se quando* *(*ele₁*) *ganhou o premio.*
the plan of Jose realised-itself when he won the prize
'Jose's plan came true when he won the prize.'
(Brazilian Portuguese, Holmberg & Sheehan 2010: 138)

On the other hand, there is an interesting similarity between the two types of control in the sense that in the standard local configurations they are both sensitive to the type of control verb in the main clause. As widely discussed in Landau (2000, 2006), the type of obligatory control of PRO, namely subject control vs object control, is largely determined by the verb type; thus, for instance, directive verbs (*tell*, *ask*, *convince*) force strict object control, while commissive verbs (*promise*, *pledge*, *vow*, etc.) typically require subject control. This distinction implies that the narrow syntactic derivation of the relation of control needs to allow for the seemingly non-local relation of subject control across the object in the case of commissives. Now, control of *pro* appears to be sensitive to the distinction in the control verb type and it also shows for the subject control of *pro* across the intervening object:

(34) *O João   prometeu   à Maria   que   iria   embora.*
      the João   promised   to.the Maria that would go   away
      'João promised Maria that he would leave.'
      (Brazilian Portuguese, Holmberg & Sheehan 2010: 143)

The two types of control are very different in the way they can or cannot ignore syntactic islands (Ross 1967). Obligatory control of PRO cannot hold across islands, with the *wh*-island constituting an exception, whereas the control of *pro* seems to be oblivious to island domains and holds across them freely, for example across a boundary of the clausal adverbial/adjunct in (35) and into a noun complement clause in (36):

(35) *Eeva saa tulla   muukan jos (hän) lupaa   olla hiljaa.*
      Ewa   may come along   if   she   promises be   quiet
      'Ewa may come along if she promises to be quiet.'
      (Finnish, Holmberg & Sheehan 2010: 147)

(36) *O presidente$_1$ negou   os rumores   de que (elle$_1$) tinha recebido*
      the president   denied   the rumours of that he   had   received
      dinheiro de   empresários.
      money   from   businessmen
      'The president denied the rumours that he had received money from businessmen.'
      (Brazilian Portuguese, Holmberg & Sheehan 2010: 148)

Properties of the control of *pro* and PRO are summarised below in (37) (Holmberg & Sheehan 2010: 141).

(37)

	Brazilian Portuguese	Finnish	Marathi	English obligatory control PRO
Binder must be [+human]	No	No	No	No
Independent time reference	Yes	Yes	Yes	No
Next clause up condition	No	No	No	Yes, otherwise it's non-obligatory control
Binder must c-command *pro*	Yes	No control under connectivity	Yes	No-control under connectivity
Split antecedent	Yes	Yes	Yes	No
Sloppy reading only with ellipsis	Yes	No	No	Yes

Although Holmberg & Sheehan comprehensively and extensively present key similarities and differences between this type of control and obligatory control of PRO, they fall short of providing a formal definition of their concept of this relation. The closest they come to it is when they propose delayed Spell-out to allow for the control relation to hold within the same derivational phase (Holmberg & Sheehan 2010: 151–152):

(38) $[_{CP3} \ldots DP \ldots [_{CP2} \ldots [_{CP1} [uD] \, C \, TP \,]]]$

This representation pertains to example (31) discussed above, where the controller DP is placed a clause further removed from the clause containing the null subject. Holmberg and Sheehan assume that the null subject moves to the edge of CP to be accessible to its controller from the outside of the phase, in line with the Phase Impenetrability Condition of Chomsky (2001).[8]

    This is, however, insufficient, as the CP_2 phase separates them. Thus, they are led to acknowledge that "the spell-out of the uD-marked pronoun at the edge of

---

8 The Phase-Impenetrability Condition (PIC):

(i)   The domain of H is not accessible to operations at ZP (with ZP the smallest strong phase), only H and its edge are accessible to such operations (Chomsky 2001: 14).

(ii)   Interpretation/evaluation of phase α takes place uniformly at the next higher phrase, i.e. Ph_1 is interpreted/evaluated at the next relevant phase Ph_2 (Chomsky 2001: 13).

(iii) $[_{ZP} Z \ldots [_{HP} α \, [H \, YP]]]$ with ZP and HP as strong phases (Chomsky 2001: 14).

CP$_1$ in (38) can wait until a DP is merged in CP$_3$, provided that there is no c-commanding DP in CP$_2$" (Holmberg & Sheehan 2010: 151).

Thus, the critical elements of the relation of *pro* control include: (a) c-command by the controller DP; (b) movement of the [*u*D] pronoun to Spec,CP to meet the strictures of the PIC; (c) some form of phase sliding (Gallego 2010) or phase extension (Den Dikken 2007), as defined above.

3.2 Articulated left periphery approach

As mentioned above in the context of (23)-(24), current discussion of the licensing and interpretation of *pro* in consistent NSL includes the notion of the Aboutness-Shift Topic. It correlates formal grammar with discourse properties and identifies a referential *pro*. This Topic is defined in the following manner:

(39)   *Topic Criterion* (Frascarelli 2018: 212)
    a.   The high Topic field in the C-domain contains a position in which the [+aboutness] feature (an extended EPP feature) is encoded and matched (via Agree) by the local (third person) null subject.
    b.   When continuous, the [+aboutness] topic can be null (i.e. silent).

This type of Topic (the A-Topic) occupies the position in ShiftP, while the familiar/given Topic occupies a lower position of the Familiar Phrase in the articulated left periphery:

(40)   [$_{ForceP}$ [$_{ShiftP}$ [$_{GP}$ [$_{ContrP}$ [$_{FocP}$ [$_{FamP}$ [$_{FinP}$]]]]]]]

The A-topic is tracked through the discourse and within complex sentences. The tracking relies on the following conditions:

(41)   *The Topic Chain Condition* (Frascarelli 2017: 217–222)
    a.   An A-Topic chain can only be started from a root (or root-like C-domain).
    b.   The [aboutness] feature is maintained through a silent G-Topic in the embedded C-domain, or
    c.   It is maintained through an overt G-Topic, such as a low-toned DP or pronoun.

The application of conditions (41b-c) is illustrated with Italian examples below:

(42) [$_{ShiftP}$ *Maria*$_k$ [$_{AgrsP}$ pro$_k$ *pensa* [$_{ForceP}$ *che* [$_{FamP}$ <*Maria*$_k$> [$_{IP}$ pro$_k$ *vincerà la gara*]]]]]
      'Maria$_k$ thinks that (she$_k$) will win the race.'

(43) [$_{ShiftP}$ *Maria*$_k$ [$_{AgrsP}$ pro$_k$ *pensa* [$_{ForceP}$ *che* [$_{FamP}$ *lei*$_k$ [$_{IP}$ pro$_k$ *vincerà la gara*]]]]]
      'Maria$_k$ thinks that she$_k$ will win the race.'

Significantly, regular G-Topics are not supposed to interfere with A-Topic chains. The examples above indicate that Frascarelli assumes that overt and null subjects in consistent NSLs occupy distinct syntactic positions. While the former are placed in C-domain of the left periphery (here FamP), the latter occupy Spec,TP. This idea is not new, as the proposal that pre-verbal lexical subjects in Romance languages occupied an A-bar position was already in circulation in the GB era. The use of a strong accented subject pronoun in (43) would be highly unnatural, which Frascarelli (2018: 219) captures through the Avoid Pronoun Principle, harking back to Chomsky (1981):

(44) *Avoid Pronoun Principle*
      Avoid strong pronoun, whenever it agrees with the current A-Topic.

The A-Topic related account centres on 3rd person pronouns, because it appears that both 1st and 2nd person null subjects do not require the formation of the A-Topic chain and are licensed through alternative means. Frascarelli captures this fact and incorporates it into her analysis by adopting the position of Sigurðsson (2011) holding that the C-domain houses syntactically active speaker and hearer features, known as the "logophoric agent (Λ_A)" and "logophoric patient (Λ_P)". These features license 1st and 2nd person *pro*. Together with the A-Topic, they belong to the same class of licensors of *pro*, the so-called C/Edge linkers (CLns):

(45) C/*Edge-Linking Generalisation*
      Any definite argument, overt or silent, positively matches at least
      one CLn in its local CP-domain, CLn ∈ { Λ_A, Λ_P, Top . . .}

Frascarelli shows that the C/Edge linkers do not interfere with one another, specifically, the formation of the A-Topic chain is not broken up by appearance of 1st or 2nd person null subjects.

At the end of the previous section certain problems with formulating a formal definition of the control of *pro* in partial NSLs are mentioned. They can be avoided to some extent, if one follows an idea outlined in Frascarelli (2018), where it is proposed that the A-Topic analysis of *pro* can be extended to partial NSLs as well.

Frascarelli presents results of her carefully designed experiments showing that the differences in the licensing of the null subject in the embedded clause of all types and designs in Italian (a consistent NSL) and Finnish (a partial NSL) are not as big as expected on the assumption that both languages should apply very distinct strategies. For example, judgements elicited for the adverbial clause of condition are quite similar (Frascarelli 2018: 228–229):

(46) a. *Leo   può venire   se pro   finisce   il lavoro.*
         Leo   can come.INF if (he) finish.3SG   the work
         'Leo can come if/when he finishes work.'
         (Italian, Frascarelli 2018: 228)

     b. *Leo   voi tul-la   jos pro   sa-a   työ-n   teh työ.*
         Leo .NOM can come.INF if   (s/he) get.3SG work.ACC do.PST.PTCP
         'Leo can come if (s/he) finishes the work.'
         (Finnish, Fracarelli 2018: 228)

(47) Who is *pro?* Leo   somebody else both (ambiguous)
      Italian   66%   12%   22%
      Finnish   75%   7%   18%

For Holmberg & Sheehan (2010), the positive evaluation of this construction stems from the fact that the main clause subject c-commands the embedded adverbial clause with *pro* in it. Now, the judgements of the Finnish example should be very different once the order of the main and embedded causes is reversed, given that here *pro* in the adverbial clause has no c-commanding antecedent. Surprisingly, for quite a few experiment participants the null subject could still be licensed:

(48) a. *Se pro finisce   il lavoro, Leo   può venire.*
         if   finish.3SG the work Leo   can come.INF
         'If (he or she) finishes the work, Leo can come.'   (100% OK)
         (Italian, Frascarelli 2018: 229)

     b. *Jos pro teke-e   työ-n,   loppu-un, Leo   voi tul-la.*
         if   make.3SG work.ACC end.ILL   Leo.NOM can come.INF
         'If (he or she) finishes the work, Leo can come.'   (48% OK)
         (Finnish, Fracarelli 2018: 229)

(49) Who is *pro?* Leo   somebody else both (ambiguous)
      Italian   45%   23%   32%
      Finnish   26%   48%   26%

Frascarelli takes these facts to indicate that the A-Topic chain approach can be extended to partial NSLs. However, it is overt (rather than covert) links in the A-Topic chain that facilitate its interpretation. In other words, the interpretation of the null subject in the embedded clause does not require a c-commanding DP antecedent but the interpretation of the chain is more prominent if a lexical pronoun placed in the C-domain overtly marks a link in the A-Topic chain. Crucially, overt links in the A-Topic chain facilitate the interpretation of *pro*, while covert links make this interpretation more difficult but not impossible.

Frascarelli (2018: 238–239) proposes to capture this distinction between consistent and partial NSLs through a mesoparameter:

(50) *Mesoparameter: The Interface Visibility Condition* (Frascarelli 2018: 238–239)
In partial NSLs:
(a) the [+aboutness] feature is connected with a P(honological)-feature in Shift [see (40)], satisfying the interface requirement that (at least) one link of the topic chain be visible at the interface levels;
(b) Minimal (and semantically eligible) overt links optimise the interpretation of topic chains at the (PF, LF) interfaces.

The extension of the A-Topic approach to the licensing of embedded null subjects in partial NSLs is explored in a few contributions to this volume (see, e.g., Madariaga's discussion of Russian).

3.3 Rootless nP approach

The two major accounts discussed above, which have become genuinely intertwined, do not fully exhaust the menu of possibilities outlined in current syntactic literature. Barbosa (2019) draws a close parallel between the possibility of licensing null NP arguments and *pro*, specifically the *pro* found in partial and radical NSLs (Barbosa refers to them as discourse NSLs). She takes such *pro* to constitute an instantiation of a rootless nP, a minimal nominal projection without the content root corresponding to the NP complement of n.

Barbosa observes that the account of null subjects in consistent and partial NSLs proposed by Holmberg encounters a challenge in examples of the following type:

(51) *Estão   a   bater   à   porta*.
　　　are   at knock.INF at.the door
　　　'There is someone knocking at the door.'
　　　(European Portuguese, Barbosa 2019: 492)

This European Portuguese example is ambiguous: the 3rd person null subject can either be interpreted as referring to a specific, contextually defined group of people (definite reading) or have indefinite reading, equivalent to *someone*, excluding the speaker and the addressee. So apparently, with the latter reading, there is no (null) aboutness shift topic in this structure and it is not clear how the [*u*D] feature of T is valued. She further argues that positing a null locative element that could check the EPP property of T is counter-productive. Although it could account for why (51) is grammatical, it would lead to overgeneration in the case of consistent NSLs, because they would also be expected to allow for an inclusive indefinite reading of indefinite *pro*.⁹ Instead, Barbosa draws a parallel between the null subjects in embedded clauses in partial and radical NSLs. It turns out that these parallels are close, specifically both types of *pro* share the same restriction on distinct intervening subjects. Both types of NSLs also share the capacity to license bare NP arguments and null objects. She follows Tomioka (2003), who proposes that in languages whose nominal phrases do not require determiners (non-DP, or NP languages in Bošković 2012), bare NPs and null NPs share basic semantic properties. The bare NP in Japanese is basically of the semantic type $\langle e,t \rangle$ anaphora, and the grammar of Japanese applies two semantic operations to derive its two prominent interpretations:

(52) *Existential Closure* (Heim 1982): ∃-closure
　　　For any $P \in D \langle e,t \rangle$
　　　∃-closure $(P) = \exists x.P(x)$
　　　(cited in Barbosa 2019: 495)

(53) *Type-shifting of a predicate to an individual* (ι) (Partee 1987): *Iota*
　　　For any $x \in D, P \in D \langle e,t \rangle$
　　　$\iota(P) = \iota x.P(x) =$ the unique x such that $P(x)$
　　　(cited in Barbosa 2019: 495)

---

9 In Holmberg's analysis this reading is blocked because the topic valuing [D] on T also satisfies EPP of T; lack of the topic should violate EPP and cause ungrammaticality but this is exactly what happens in (51). There is no topic, [*u*D] of T is apparently not satisfied, but the sentence is fine.

Translated into the language of syntax, the definitions in (52) and (53) mean that when interpreted in the vP internal position, a bare noun and the null subject in radical and partial NSLs is interpreted as indefinite or impersonal in the domain of existential closure (see also Tsedryk 2022: 49–57). When the operation of individuation applies, and possibly the bare NP or *pro* occupies a syntactic position outside vP, they can be interpreted as definite, see example (56) below for illustration. Tomioka claims that the null subject in Japanese is a null NP "whose descriptive content is pragmatically retrieved" (Barbosa 2019: 495). This procedure of retrieval is similar to the analysis of elliptical structures involving null NPs, such as:

(54) *John bough one book, I bought five* [$_{NP}$ ___].

Barbosa applies Tomioka's observations also to DP languages and notices that in certain DP languages the determiner can be null and still license null NPs, while in others (English or German) only lexical determiners can license null NPs. Following Panagiotidis (2014), Dvořák (2015) and Ruda (2017), she proposes to treat *pro* as a null nP without a root (content):

(55) [$_{nP}$ n]

When such nP is selected by a null D, the argument still remains null, but when it is selected by a lexicalized D, a lexical pronoun obtains. Barbosa proposes to interpret *pro* in partial and radical NSLs as resulting from, and retrieved through, mechanisms analogous to the NP ellipsis/anaphora, as in (54). Following Vainikka & Levy (1999), she observes that in Finnish two distinct readings of the null subject, consistent with the definitions in (52) and (53), correlate with distinct syntactic positions occupied by *pro*:

(56) a. *Oppilas   tietää   ettei   tehtävää   pysty ratkaisemaan.*
          student   knows   that.not   assignment   can   solve
          'The student knows that the assignment can't be solved.'
          not 'The student knows that he can't solve the assignment.'
          (Finnish, Barbosa 2019: 502)
      b. *Oppilas   tietää   ettei   ___   pysty ratkaisemaan   tehtävää.*
          student   knows   that.not   can   solve   assignment
          'The student knows that he can't solve the assignment.'
          Not 'The student knows that the assignment can't be solved.'
          (Finnish, Barbosa 2019: 502)

In (56a), the object of the embedded clause satisfies the EPP and *pro* obtains the impersonal indefinite reading; it also probably remains in Spec,vP. In (56b), on the other hand, where apparently *pro* checks the EPP property of T and occupies Spec,TP, the definite interpretation of *pro* is available. Barbosa submits that such syntactic and morphosyntactic operations as topicalization, movement to the pre-verbal position in the clause, rich φ-feature agreement and specifically the marking for [person], facilitate the application of the semantic mechanism of individuation in (54) and the definite interpretation of *pro*. The interpretation of *pro* in these languages involving ellipsis or anaphora makes it analogous to nominative anaphors in Chinese; after all, its interpretation hinges on an antecedent and it occupies a position where nominative case is licensed. Interestingly, the lexical nominative (subject) anaphor *ta ziji* in Chinese shares the same locality restrictions with *pro* in radical and partial NSLs, see (57); an intervening distinct subject disallows the binding of *ta ziji* by a more remote antecedent:

(57) *Xiaoming₁ shuo Zhangsan₂ xiangxin ta ziji*₁/₂ neng kaoguo.*
      Xiaoming say Zhangsan believe him self can pass.the.exam
      'Xiaoming₁ says that Zhangsan₂ believes that he*₁/₂ can pass the exam.'
      (Chinese, Barbosa 2019: 514)

This fact leads Barbosa to conclude that this type of null subject is a defective nP, lacking φ-features, like anaphors, and it denotes an individual in the domain.

Barbosa proposes to extend her account of *pro* as nP in partial NSLs also to consistent NSLs. She follows Holmberg (2005) and assumes that here the head T bears both a set of interpretable φ-features and the D-feature:

(58) $[_{TP} [_T v_i - T{<}D_k\varphi{>}][_{vP} nP [_{v'} t_i \ldots]]]$

Null nP denotes a property (of type $\langle e,t \rangle$) and although it may seem incompatible with vP (also of the semantic type $\langle e,t \rangle$), it can be combined with it through a semantic operation of Predicate Modification.[10] The index provided by feature D in T applies to the complex property resulting from the instance of Unification. If this proposal is accepted, *pro* in both partial and consistent NSLs can be taken to be a minimal argument, the indefinite minimal nP placed in Spec,vP.

---

10 Cf. Farkas & de Swart's (2003) Unification. For more details, see also Tsedryk (this volume).

3.4 Other approaches

Another question connected with *pro* is put on the research agenda in Kučerová (2014). The author focuses on an interesting observation pertinent to the positioning of *pro* in the syntax. Despite the lore established in the field, Kučerová provides interesting arguments to the effect that the position of *pro* in consistent NSLs does not need to be identified with Spec,TP, which is typical of Spanish, Italian and Greek. She proposes that in a subset of consistent NSLs the *pro* subject occupies a lower position, namely the thematic position of Spec,vP. The reason for this state of affairs is that in these languages *pro* fails to satisfy the EPP and one of the key diagnostics is a ban on verb first word orders in main clauses in a selected set of constructions. Kučerová shows that Czech, Old French and Modern Hebrew belong to this subset of consistent NSLs with *pro* placed in Spec,vP. Kučerová (2014: 134) formulates the following prediction capturing the close relationship between pro and the EPP requirement:

(59) *The D-feature Hypothesi*s (Kučerová 2014: 134)
     If a null subject language does not have an overt element in the specifier
     of the inflectional projection, the EPP requirement is satisfied either by *pro*
     being merged in this position or by a D-feature being in T.[11]

This prediction works perfectly for Italian in (61a) but, surprisingly, does not seem to work perfectly for Czech in (60b), among others:[12]

(60) a. *È   arrivato.*
         is   arrived
         'He arrived.'
         (Italian, Kučerová 2014: 133)
     b. **Jsem   přišel   domů   pozdě.*
         AUX.1SG   come.PTCP   home   late
         'I came home late.'
         (Czech, Kučerová 2014: 135)

---

11 The reference to the D-feature in T is related to the proposal in Alexiadou & Anagnostopoulou (1998) holding that the EPP can be satisfied by the Tense head itself, provided it houses the D-feature (pronominal features).
12 The standard reason for ungrammaticality of such constructions in Czech grammar is taken to be the clitic status of the auxiliary verb in the verb first position (e.g., Franks and King 2000; Bošković 2001). Kučerová (2014) argues against this point extensively.

Yet, this ban on verb first word orders in Czech does not hold of embedded clauses or main clauses with clear V-to-T movement properties:

(61) *Petr   řekl   mamince,   že   jsem   přišel   domů   pozdě.*
      Petr.NOM said   mother.DAT that AUX.1SG come.PTCP home   late
      'Peter told his mother that I came home late.'
      (Czech, Kučerová 2014: 136)

Kučerová proposes that these facts are better captured by an extended notion of the EPP and disassociation of the presence of *pro* in a particular grammar from its EPP-related function. In certain consistent null subject grammars *pro* satisfies the EPP (e.g. Italian, Spanish, Greek), while in others it does not (e.g., Czech, Modern Hebrew, Old French). As for the EPP, she proposes that it should be redefined as a principle forcing further merger of some material to the projection of T, once T has been merged into the syntactic object under construction:

(62) *T-Extension Requirement (TER)*
      If Merge (T, α) applies, Merge (T', β) must be the next step of the derivation, where T' is a projection of T and β belongs to the same phase domain as T.

The TER forces creation of a specifier via internal or external merge, addition of an adjunct to TP or head movement of a verbal category to T. Given that, technically speaking, TER does not require extension of the projection of T in the sense of Chomsky (1995), even an operation embedding TP in the projection of C (external merger of the complementiser) will yield a grammatical output. Hence, (61) is fine, and (60b) is not.

   Thus, if the grammar of a given consistent NSL allows *pro* to satisfy TER, the verb first word order in main clauses is always acceptable, see (61a). However, if the grammar of a consistent NSL does not allow *pro* to satisfy the TER, verb-first word orders in main clauses will be limited only to cases where a (possibly) silent T forces V-movement to it. Otherwise, such word orders should be impossible in main clauses, but clearly attested in embedded clauses, where the TER is satisfied.

   Kučerová presents another empirical argument in support of her view that *pro* may be placed in Spec,vP in some languages based on possessive pronominal binding in Czech. The nominative subject cannot bind a pronominal possessive across an intervening dative experiencer in the raising construction, see (63). Yet, once it is placed in the infinitive and the intervention effect is naturally nullified, the binding is fine, consider (64). Now, the *pro* subject as antecedent to the pronominal possessive is completely insensitive to intervention of the dative experiencer, as in (65), which leads Kučerová to conclude that *pro* binds from the

same position as the lexical antecedent in (64), that is probably Spec,vP, rather than from Spec,TP as the lexical antecedent in (63).

(63) *Petr$_i$*   *se*   *zdál*   *Marii*   *být*   *závislý*   *na*   *jeho$_{*i/j}$*
Petr.NOM   RFL   seemed   Marie.DAT   be.INF   dependent   on   his
rodině.
family
'Petr$_i$ seemed to Maria to be dependent on his$_i$ family.'
(Czech, Kučerová 2014: 141)

(64) *Včera*   *zdál*   *se*   *Marii*   *Petr$_i$*   *být*   *závislý*
yesterday   seemed   RFL   Marie.DAT   Petr.NOM   be.INF   dependent
na   *jeho$_i$*   *rodině.*
on   his   family
'Yesterday Petr$_i$ seemed to Maria to be dependent on his$_i$ family.'
(Czech, Kučerová 2014: 143)

(65) *Zdál*   *se*   *Marii*   *být*   *závislý*   *na*   *jeho*   *rodině.*
seemed.M   RFL   Marie.DAT   be. INF   dependent   on   his   family
'He$_i$ seemed to Maria to be dependent on his$_i$ family.'
(Czech, Kučerová 2014: 142)

The proposal divorcing *pro* from EPP (Kučerová's TER), with concomitant consequences for expected verb first word orders in main clauses, opens yet another avenue for research into null subjects. Seen from this angle, even consistent NSLs do not form a homogeneous group. This may be bad news for the original Null Subject Parameter in (1) and (8) but it certainly is a step forward in pursuit of empirical adequacy.

3.5 The φP, nP and A-Topic as composite parts of the GB-era *pro*?

The central idea of Holmberg's account corresponds to the view on the function of agreement morphology expressed in Rizzi (1982, 1986) and Jaeggli & Safir (1989): a sufficiently rich paradigm of inflectional morphology is indirectly crucial for the formal licensing of the null subject in a consistent NSL. The procedure of φP incorporation is contingent on a relatively rich specification of T (T/AGR), as incorporation can take place only if the features of the incorporee form a proper subset of the features of the incorporation host. While it is clear that the nominal

argument bears φ-features (person, number, gender), it can incorporate only into T which also bears and distinguishes these features. However, there is a minimalist twist to the relevance of rich agreement morphology on the verb. As Holmberg (2005: 560) underscores, rich morphology of AGR is irrelevant to narrow syntax, as narrow syntax does not see PF realisations of the φ-features, only the φ-features themselves. So rich morphology, visible at PF, contributes to sentence processing, where the recovery of the null subject is performed. For such recovery, all types of information available from the discourse count the same: both rich morphology (in consistent and partial NSLs) as well as discourse prominence (radical NSLs).

Holmberg's (2005, 2010a,b) account, supported by Frascarelli's (2007, 2018) findings, has proved to be seminal and has generated substantial amount of research. The contributions in this volume provide and test additional new data, and critically refer to the central notions of the foundation accounts. Undoubtedly, what Rizzi (1982, 1986) did to the GB era discussion of the null subject, Holmberg has done to its minimalist rendition.

Still, a few questions remain open, though. Let me just mention two, partly taken up by the authors of the articles in this collection. They both concern narrow-syntactic mechanisms used in the comprehensive description of NSLs. One issue pertains to the technical ways in which the D-feature (index) of the A-Topic is passed onto the D-feature of T in consistent NSLs. In minimalism feature sharing, in a general sense, not the technical sense of Pesetsky & Torrego (2007), requires a probe/goal relation, called Agree. Yet, which is which in the matching of the A-Topic and T? Is T a probe? It should be, as probes are heads and they usually bear unvalued features. But at the same time, the probe should c-command the goal and this is not the case, when the null A-Topic occupies a position in the CP domain and T is just too low in the structure. So can a goal c-command the probe? Several authors have proposed that a head could Agree (search for a goal) upwards (Zeijlstra 2012; Bjorkman & Zeijlstra 2019) but this issue is still a hotly debated option (Preminger 2013; Preminger & Polinsky 2015).

The second question concerns the notion of control, as defined by Holmberg & Sheehan (2010). The authors insist that this relation must be distinct from regular obligatory control of the PRO subject of the infinitive (Landau 2000, 2006; Hornstein 1999, 2001) and it must involve a direct relation between the lexical DP antecedent and the null subject, with the former c-commanding the latter. The problem seems to be that every current minimalist theory of control applies the notion of Agree or Move and these notions are subject to strict locality conditions, for which plain c-command does not suffice (for further discussion, see also Vinka, this volume). For instance, they are largely sensitive to the Phase Impenetrability Condition (PIC) of Chomsky (2001), as well as island constraints. The relation of "finite control" appears to be able to cross phase and island bound-

aries, thus it eludes strict boundaries so characteristic of most narrow syntactic phenomena. Having said that, "control-like" properties of the null subjects in embedded clauses in partial NSLs are a fact, so our current state of knowledge only whets one's appetite for more comprehensive accounts and spurs on research in this area (see Madariaga, this volume). A similar state of affairs holds of the role of the A-Topic in the analysis of the null subject in consistent NSLs. The facts concerning topic shift shown in example (23) are unquestionable but the problem is that formally restricting and systematically defining relations between discourse management and narrow syntax is a challenging task. Insightful results of Frascarelli's studies, where discourse properties are partly harnessed into the strictures of sentence grammar, compete with a less disciplined view on how the referent of *pro,* perceived as the minimal nP, can be identified.

A set of contributions that follows is supposed to broaden the empirical and theoretical picture of NSLs.

References

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrising AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539.

Barbosa, Pilar 2019. *Pro* as a minimal nP: A unified account of *pro*-drop. *Linguistic Inquiry* 50 (3). 487–526.

Biberauer, Theresa. 2018. *Pro*-drop and emergent parameter hierarchies. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar*, 94–134. Oxford: Oxford University Press.

Bjorkman, Bronwyn & Hedde Zeijlstra. 2019. Checking up on (φ-)Agree. *Linguistic Inquiry* 50 (3). 527–569.

Borer, Hagit. 1986. I-subjects. *Linguistic Inquiry* 17. 375–416.

Borer, Hagit. 1989. Anaphoric AGR. In Oswaldo Jaeggli & Kenneth Safir (eds.), *The Null Subject Parameter*, 69–109. Dordrecht: Kluwer.

Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier.

Bošković, Željko. 2012. Phases in NPs and DPs. In Ángel J. Gallego (ed.), *Phases: Developing the Framework*, 343–383. Berlin: Mouton de Gruyer.

Cabredo-Hofherr, Patricia 2003. Arbitrary readings of third person plural pronominals. In Matthias Weisgerber (ed.), *Proceedings of the conference "sub7-Sinn und Bedeutung"*, 81–94. University of Konstanz.

Cardinaletti Anna & Michal Starke. 1999. The typology of structural deficiency: a case study of the three classes of pronouns. In Riemsdijk, van, Henk (ed.), *Clitics in the languages of Europe*, 145–233. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries. In Roger Martin, David Michaels, Juan Uriagereka & Samuel J. Keyser (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik,* 89–156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Cinque Guglielmo. 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19. 521–581.

D'Alessandro, Roberta. 2007. *Impersonal SI Constructions. Agreement and Interpretation*. Berlin: Mouton de Gruyter.

D'Alessandro, Roberta & Artemins, Alexiadou. 2003. Inclusive and exclusive impersonal pronouns. A feature-geometrical analysis. *Rivista di Grammatica Generativa* 27. 34–44.

Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33. 409–442.

Den Dikken, Marcel. 2007. Phase extension: Contours of the theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33. 1–41.

Dvořák, Věra. 2015. Syntax of generic null objects revisited. In Anna E. Jurgensen, Hannnah Sande, Spencer Lamoureux, Kenny Baclawski & Alison Zerbe (eds.), *Proceedings of the Forty First Annual Meeting of the Berkeley Linguistic Society*, 71–89. Berkeley: University of California, Berkeley Linguistic Society.

Farkas, Donka & Henriette de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford, CA: CSLI Publications.

Fenger, Paula. 2018. How impersonal does *one* get? J*ournal of Comparative Germanic Linguistics*. 21 (3). 291–325.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistics Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of pro in consistent and partial null-subject languages. A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar*. 211–239. Oxford: Oxford University Press.

Franks, Steven & Tracy Holloway King. 2000. *A Handbook of Slavic Clitics*. New York: Cambridge University Press.

Gallego, Angel. 2010. *Phase theory*. Amsterdam: John Benjamins.

Harbour, Daniel, David Adger & Susane Bejar. 2008. *Phi Theory. Phi Features across Modules and Interfaces*. Oxford: Oxford University Press.

Heim, Irene. 1982. The semantics of definite and indefinite NPs. Amherst, MA: University of Massachusetts PhD dissertation.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010a. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. The null generic subject pronoun in Finnish: a case of incorporation. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric variation: Null Subjects in Minimalist Theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders & Michelle Sheehan. 2010. Control into finite clauses in partial null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–152. Cambridge: Cambridge University Press

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30. 69–96.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Huang, James 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Jaeggli, Osvaldo & Kenneth Safir. 1989. The null subject parameter and parametric theory. In Osvaldo Jaeggli & Kenneth Safir (eds.), *The null subject parameter*, 1–44. Dordrecht: Kluwer.

Kučerová, Ivona. 2014. The syntax of null subjects. *Syntax* 17 (2). 132–167.

Landau, Idan. 2000. *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer Academic Publishers.

Landau, Idan. 2006. Severing the distribution of PRO from case. *Syntax* 9 (2). 153–170.

Longobardi, Giuseppe. 1994. Reference and proper names. *Linguistic Inquiry* 25. 609–665.

Madariaga, Nerea. 2022. Referential null subjects in Russian: a synchronic and diachronic overview. In Grete Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure, and Typology*, 000–000. Berlin: Mouton de Gruyer.

Newmeyer, Frederick. 2004. Against a parameter setting approach to language variation. In Pierre Pica, Johan Rooryck & Jeroen van Craenenbroeck (eds.), *Language Variation Yearbook*, *vol. 4*, 181–234. Amsterdam: John Benjamins.

Panagiotidis, Phoevos. 2014. *Categorial Features: A Generative Theory of Word Class Categories*. Cambridge: Cambridge University Press.

Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk, Dickde Jongh & Martin Stokhof (eds.), *Studies in Discourse Representation Theory and the Theory of Generalised Quantifiers*, 115–143. Dordrecht: Foris.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation,* 262–294. Linguistik Aktuell / Linguistics Today 101. Amsterdam: John Benjamins.

Preminger, Omer. 2013. That's not how you agree: A reply to Zeijlstra. *The Linguistic Review* 30. 491–500.

Preminger, Omer & Maria Polinsky. 2015. Agreement and semantic concord: A spurious unification. Ms., Harvard University.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–557.

Rodrigues, Cilene. 2004. *Impoverished Morphology and A-movement out of Case Domains*. PhD dissertation. University of Maryland, Maryland.

Ross, John. 1967. *Constraints on variables in syntax*. PhD dissertation. MIT. Cambridge, MA.

Roberts, Ian. 2010. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 58–87. Cambridge: Cambridge University Press.

Roberts, Ian & Anders Holmberg. 2010. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Ruda, Marta. 2017. *On the Syntax of Missing Objects: A Study with Special Reference to English, Polish, and Hungarian*. Amsterdam: John Benjamins.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. PhD dissertation. Rutgers University, New Jersey.

Sigurðsson, Halldor A. 2011. Conditions on argument drop. *Linguistic Inquiry* 42 (2). 267–304.

Taraldsen, Knut Tarald. 1980. *On the nominative island constraint, vacuous application and the that-trace filter*. Indiana University Linguistics Circle. Bloomington: Indiana.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its crosslinguistic implications. In Kerstin Schwabe and Susanne Winkler (eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*, 321–340. Amsterdam: John Benjamins.

Tsedryk, Egor. 2022. Null subjects and their overt counterparts in East Slavic root clauses: Referential and non-referential readings. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure, and Typology*, 000–000. Berlin: Mouton de Gruyter.

Vainikka, Anne and Yonata Levy. 1999. Empty subjects in Hebrew and Finnish. *Natural Language and Linguistic Theory* 17. 613–671.

Vinka, Mikael. 2022. Two types of null subjects in South Saami. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure, and Typology*, 000–000. Berlin: Mouton de Gruyter.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29. 491–539.

Egor Tsedryk

2 Null subjects and their overt counterparts in East Slavic root clauses: Referential and non-referential readings

1 Introduction

Partial Null Subject Languages (NSLs) have two definitional properties: (i) the availability of a generic NS, which may appear in main and dependent clauses, and (ii) the availability of an anaphoric NS, found in syntactically embedded contexts and dependent on a matrix antecedent (Holmberg 2005, 2010b; Holmberg, Nayudu & Sheehan 2009; Holmberg & Sheehan 2010). This chapter deals with the first type of NS in East Slavic languages, including Belarusian, Russian and Ukrainian (henceforth BRU).[1] While concentrating on root clauses, this chapter investigates pronominal subjects in BRU at the interfaces with respect to their form (overt or null) and interpretation (referential or non-referential).[2]

Analyzing the data from Brazilian Portuguese, Marathi and Finnish, Holmberg, Nayudu & Sheehan (2009) and Holmberg (2010b) claim that the crucial difference between consistent and partial NSLs lies in the distribution of the third person plural (3PL) NS, equivalent to 'they' ('In Italy they drink espresso'), and the third person singular (3SG) NS, corresponding to 'one' ('One can sit comfortably in this chair'). Thus, according to these authors, consistent NSLs have the former but lack the latter, and, in order to obtain the 'one' reading, these languages have to resort to alternative strategies, including the generic 2SG and the reflexive *se* (see Holmberg, Nayudu & Sheehan 2009: 63–64). These strategies are also used in BRU, which do not have a generic 3SG subject either, as shown in (1).[3]

---

1 The East Slavic languages form a homogenous group with respect to the NS properties but Russian received the greatest attention in the literature. Many authors agree that Russian is a partial NSL (Tsedryk 2015; Bizzari 2015; Frascarelli 2018; Madariaga 2018; Barbosa 2019). Madariaga (2022: 75–104) gives a more extensive discussion, focusing exclusively on Russian from both the synchronic and the diachronic perspectives.

2 A non-referential (arbitrary) reading arises when a variable is bound in a given domain without fixing its reference set. This type of reading can be existential or generic (quasi-universal). A generic reading arises when a variable ranges over a maximal set of referents in a given domain (e.g., a given location).

3 The examples provided in this chapter reflect my own judgements and were constructed in consultation with native speakers of the respective languages. For reasons of space, I will not

(1)   a.   *Kavu   pjuc'   ranicaj.*
      coffee.ACC   drink.3PL   morning.INS
      'Coffee is drunk in the morning.' / 'One drinks coffee in the morning.'
      (Belarusian)

      b.   *Tut   ne   porabotaeš'.*
      here   NEG   will.work.2SG
      'It is impossible to work here.' / 'One can't (won't be able to) work here.'
      (Russian)

      c.   *Tut   pracjujet'sja   dobre.*
      here   work.3SG.RFL   well
      'It feels good to work here.' / 'One feels good working here.'
      (Ukrainian)

The NS makes it possible to formulate generic statements referring to people in
general. If an overt pronoun were used in (1), only the referential reading would
emerge, as shown in (2). Examples (2a) and (2b) feature overt nominative pro-
nouns, whose person and number features match those of the verbal inflection,
i.e. 3PL and 2SG, respectively. In (2c), on the other hand, there is no agreement
between the overt dative pronoun and the 3SG verbal inflection.

(2)   a.   *Jany   kavu   pjuc'   ranicaj.*
      they.NOM   coffee.ACC   drink.3PL   morning.INS
      'They drink coffee in the morning.'
      (Belarusian)

      b.   *Ty   tut   ne   porabotaeš'.*
      you.NOM   here   NEG   will.work.2SG
      'You won't be able to work here.'
      (Russian)

      c.   *Meni   tut   pracjujet'sja   dobre.*
      me.DAT   here   work.3SG.RFL   well
      'It feels good for me to work here.' / 'I feel good working here.'
      (Ukrainian)

---

replicate the same examples in each language (the same examples can be constructed in the
other two languages). Finally, I do not use the *pro* marker in the examples throughout the text for
two reasons: (i) to facilitate the reading of the examples without any additional symbols, which
might add a biased view of the data, and (ii) to avoid confusion with regard to *pro*'s structural po-
sition (e.g., there is no way to know whether *pro* precedes or follows the accusative object in (1a)).

Returning to the examples in (1), it should be noted that genericity in these examples is determined at sentence level and it is not the inherent property of the NS. Thus, removing the adverbial in (1a) would result in an episodic (speech-time anchored) sentence in which the NS has a referential reading. The sentence in (3a) should be read as an answer to the question: 'What are they doing?'. Similarly, the question in (3b) has a speech act participant referent, being oriented towards the addressee. Finally (3c), without -*sja* and the locative adverb, is a felicitous utterance if the referent is perceptually accessible during the speech act.

(3)   a.   *Kavu   pjuc'.*
           coffee.ACC   drink.3PL
           'They drink coffee.'
           (Belarusian)

      b.   *Tut   ne   porabotaeš'?*
           here   NEG   will.work.2SG
           'Can't you work here?'
           (Russian)

      c.   *Dobre   pracjuje.*
           well   work.3SG
           'She/he is working well.' (pointing at a working person)
           (Ukrainian)

The examples in (3) imply a contextually salient (more precisely, deictic) referent, which must be easily retrievable from the ongoing situation. More generally, the NS cannot be properly interpreted without such contextual support unless a non-referential meaning is intended.

The referential/non-referential (in parallel with the overt/null) contrast sometimes corresponds to the strong/weak dichotomy in the pronominal systems of other languages (e.g., Dutch; see Gruber 2017). Holmberg (2005) attributes the various interpretative differences (along with the null/overt dichotomy) to the syntactic complexity of pronominal elements (based on earlier proposals in Cardinaletti & Starke 1999 and Déchaine & Wiltschko 2002). More precisely, a non-referential pronoun is a phrasal category containing φ-features without a determiner (φP), whereas a referential pronoun is a full Determiner Phrase (DP). Barbosa (2019) has recently proposed that a NS is just a bare nominal head, which has a phrasal status (nP).[4] This head/phrase is of type ⟨*e,t*⟩ (property) and

---

4 Little n is just a categorial head without a lexical root (see Section 4.1).

can undergo a type shift to ⟨*e*⟩ (individual). This chapter adopts Barbosa's analysis, while maintaining the φP/DP dichotomy assumed in the previous work.

The BRU pronominal system will be analyzed in the framework that takes person features as functions operating on a set of possible referents (Ackema & Neeleman 2018). The underlying idea of this chapter is that adding φ-features to the bare nP does not automatically provide a referential index that would link that nP to a discourse referent. That is, φ-features do not automatically instantiate the argument of a function with respect to a discursive referent: a D-head is required for this type of instantiation (see Farkas & de Swart 2003: 35). In terms of semantic types, a φP is still a function of type ⟨*e,t*⟩ (just like the bare nP assumed in Barbosa 2019). Thus, merging D is a syntactic way to shift from type ⟨*e,t*⟩ to type ⟨*e*⟩, but it does not mean that a φP cannot undergo this type of shifting without the D-layer. The point is that shifting to type ⟨*e*⟩ does not come for free in a D-less structure and requires certain licensing conditions. One of such conditions is movement to the specifier position of Tense Phrase (Spec,TP). In addition to the semantic effect, merging D also has an effect on spell-out. That is, D conditions the overtness of a pronoun under most – but not all – circumstances. There are two cases when a φP or a bare nP can be spelled out as an overt form, i.e. the generic 2SG and *-sja*, respectively. Section 5 will use Ackema & Neeleman's (2018) framework (outlined in Section 4) to explain why 2SG has a variable behaviour.

All in all, the goal of this chapter is to give a comprehensive picture of overt and null subject pronouns in East Slavic at the two interfaces, Phonological Form (PF) and Logical Form (LF).

This chapter is organized as follows. Section 2 revisits referential subject drop in relation to information structure, namely the partition into the given/familiar part and the comment part. It is argued that this use of NS does not involve a Topic[5] operator in the left periphery of the clause.[6] The following sections then revisit the three non-referential cases outlined in (1). Section 3 is a descriptive overview of non-referential 3PL (existential and generic occurrences). Subsequently, section 4 provides an analysis of 3PL based on Barbosa's (2019) insights. This section also introduces Ackema & Neeleman's (2018) framework, preparing the ground for the analysis of 2SG in section 5.1, while section 5.2 deals with the generic use of *-sja*. Section 6 concludes the discussion.

---

5 When capitalized, Topic is understood as an entity-referring constituent (Krifka 2007: 41; see fn. 7).
6 See Madariaga (2022) for an alternative point of view.

2 Referential NS and information structure

In BRU, subjects can be dropped irrespective of the person and tense features of the verb, which agrees in person and number in non-past tenses and gender or number in the past tense. A NS is possible in any tense and person, provided the context is relevant. Consider the examples in (4), representing answers to the question 'What are you doing?' with varying person, number, gender and tense specification.

(4) a. *Pracuju   nad   praektam.*
         work.1SG   on   project
         'I am working on a/the project.'
         ['What are you doing?']

  b. *Pracue   nad   praektam.*
         work.3SG   on   project
         'He is working on a/the project.'
         ['What is she/he doing?']

  c. *Pracavaŭ   nad   praektam.*
         worked.M   on   project
         'I (male) / he was working on a/the project.'
         ['What were you / was he doing?']

  d. *Pracavala   nad   praektam.*
         worked.F   on   project
         'I (female) was / she was working on a/the project.'
         ['What were you / was she doing?']

  e. *Pracavali   nad   praektam.*
         worked.PL   on   project
         'We were / they were working on a/the project.'
         ['What were you (plural) / were they doing?']
         (Belarusian)

The above utterances provide just the right amount of information elicited in the question. The author of the question knows or assumes that the subject's referent is involved in some activity and wants to know the details. In other words, the utterances in (4) follow up on the event already presupposed in the discourse.

Consider now a similar case in (5). Assume that the utterer of (5a) already knows that something is going on with the addressee's father, but she/he wants an update on his current condition. In this context, the subject is easily dropped in (5b).

(5)   a.   *Čto   s   otcom?*
          what   with   father.INS
          'What's up with father?' (knowing that father is not well)
      b.   *Vsë   eščë   boleet.*
          all   still   is.sick.3SG
          'He is still sick.'
          (Russian)

Given the context described above, the answer in (5b) is not just a mere comment on the father. Rather, it relates to the father's health condition. That is, from the point of view of information structure, what is given/known here is the whole proposition (or a set of propositions), which forms the Common Ground (CG) content in (5). This is the implicit content that the interlocutors rely on in communication (Krifka 2007: 15).[7] In this sense, we can consider the subject of (5b) as a Given/Familiar Topic, or G-Topic (Frascarelli 2007, 2018; Frascarelli & Hinterhölzl 2007; Bianchi & Frascarelli 2010). Crucially, this Topic is part of the CG content, but it is not the sole element of this content. The subdivision into the asserted part (the actual utterance) and the implicit CG content is informally presented in (6).

(6)   [$_{\text{CG content}}$ Father is not doing well] [$_{\text{Utterance}}$ *pro* is still sick]   (see (5b))

Now suppose that the speaker does not rely on any shared information about the father and only knows that something went wrong with him. In this context, answer (7b) to question (7a) is not an update on the current state of the father's health, but new information about him. The subject cannot be dropped in this case, even if the entity has already been introduced.[8]

---

7 Literature on information structure abounds. This chapter relies on the basic notions as outlined by Krifka (2007), who takes stock of previous work, including Lambrecht's (1994) classic. More precisely, two notions are used here, i.e. Common Ground (CG) and Topic. CG has two dimensions: content and management. The first dimension is particularly relevant here, i.e. it is understood as "the information that is mutually known to be shared and continuously modified in communication" (Krifka 2007: 15). As for Topic, it is a constituent that "identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content" (Krifka 2007: 41). In this sense, Topic may or may not be part of what constitutes the theme in the theme/rheme partition of a sentence from the Prague school functional perspective (Sgall, Hajičová & Panenová 1986; see also Bailyn 2012: 257 and references therein).
8 The utterance in (7b) could possibly be grammatical in another context, but not as an answer (7a).

(7) a. *Čto-to   slučilos'   s   otcom?*
        something happened   with   father
        'Did something happen to father?'

  b. **(On)   zabolel.*
      he.NOM   got.sick.M
      'He is still sick.'
      (Russian)

As was mentioned above, the verb does not agree in person in the past tense in BRU and the lack of person agreement could potentially cause the problem in (7b). Nevertheless, the same type of infelicitous answer can be observed in (8b), where the verb does show person agreement. No previous knowledge about Ivan is assumed in (8a), except that the entity under the question ('Ivan') is part of the CG content. This is an instance of an Aboutness G-Topic (Frascarelli 2018: 215).

(8) a. *Vy   ščo-nebud'   znajete   pro   Ivana?*
        you.NOM something.ACC know.2PL about Ivan
        'Do you know anything about Ivan?'

  b. **(Vin)   zaraz pracjuje   na   zavodi.*
      he.NOM   now   work.3SG on   plant
      'He is now working at a plant.'
      (Ukrainian)

The informal representation of (8b) is given in (9), where it is shown that an entity is the only CG content. There is no propositional content that the speakers can rely on.

(9) [$_{\text{CG content}}$ Ivan] [$_{\text{Utterance}}$ *pro is now working at a plant]   (see (8b))

The examples in (7) and (8) show that being a Topic in BRU is not sufficient for *pro*-drop. The subject has to be part of the shared propositional CG content, in which case it can be easily dropped irrespectively of the agreement pattern.

Let us now observe the subject drop (or the impossibility thereof) in a sequence of sentences. In (10) – to be read as a coherent text from (10a) to (10d) – sentence (10a) introduces the Aboutness Topic (A-Topic). The A-Topic is now part of the given content, which (from this point onward) also includes a proposition (the moving event). Next, (10b) and (10c) repeat the same Topic, but do not continue the moving event. They serve to add new propositional content to the Topic, i.e. (10b) is about Ivan's mental disposition to move further and (10c) is

about his freezing in place (no movement at all). As we can observe, the pronouns cannot be dropped in these cases. However, once (10d) takes over the propositional content introduced in (10a) (continuation of the moving event), the subject is dropped.

(10) a. *Ivan*ᵢ   *vybeh*   *na*   *vulicu.*
     Ivan.NOM ran.out to   street
     'Ivan ran outside.'

    b. *(*Jamu*ᵢ) *xacelasja behčy*   *nasustrač* *soncu.*
     him.DAT wanted   run.INF   towards   sun
     'He wanted to run toward the sun.'

    c. *(*Ën*ᵢ) *zastyŭ* *na* *nejkae* *imhnenne,*
     he.NOM   froze   for   some   moment
     'He froze for a moment,'

    d. *a*   *potym* *imkliva* *pamčaŭsja* *ŭperad.*
     and   then   quickly rushed   forward
     'and then rushed forward.'
     (Belarusian)

Consider now another sequence in (11). The sentence in (11a) introduces the A-Topic and the proposition ('we are going to the lake'), which can be used as given information for the next sentence. In fact, in (11b) it is explained why we are going to the lake, relying on the propositional content of the preceding sentence. The subject is dropped, as expected. Sentence (11c), on the other hand, starts a new propositional content ('Ivan has promised to pass the keys'), and the subject cannot be null (even though 'Ivan' has already been introduced in the discourse).

(11) a. *My*   *idëm*   *na* *ozero.*
     we.NOM   are.going.1PL to   lake
     'We are going to the lake.'

    b. *Nadeemsja* *tam* *vstretit'* *Ivana*ⱼ.
     hope.1PL   there see.INF   Ivan.ACC
     'We hope to see Ivan there.'

    c. *(*On*ⱼ) *nam* *obeščal* *peredat'* *ključi.*
     he.NOM   us.DAT promised pass.INF   keys.ACC
     'He has promised us to pass the keys.'
     (Russian)

Finally, in (12) the Topic (*Ivan*) is introduced in the first sentence (12a).[9] However, there is no propositional content that the other sentence (12b) can rely on (the first sentence entails that the addressee does not know Ivan). Therefore, the subject cannot be dropped in (12b).

(12) a. *Ja   xoču   vam   predstavyty   Ivana$_i$.*
     I.NOM   want.1SG   you.DAT   introduce.INF   Ivan.ACC
     'I want to introduce Ivan to you.   '

 b. *(*Vin$_i$*)   *pracjue*   *z*   *mnoju*   *u*   *viddili*   *kadriv.*
     he.NOM   work.3SG   with   me   in   human   resources
     'He works with in human resources.'
     (Ukrainian)

The data presented in this section, complementing the cases of deictic referents in (3), lead to the following descriptive generalizations:

(13) a. The subject of a root clause can be dropped if and only if:
     (i) it is part of a propositional CG content, or
     (ii) it has a deictic referent in the shared situation.
 b. Neither A-Topics nor G-Topics can be dropped unless the above conditions are met.

It means that subject drop in East Slavic cannot be reduced to a mere Topic chain, as suggested, for example, by Frascarelli (2018) for consistent NSLs. For the sake of discussion, let us take a cartographic view of the C-domain in (14) (Frascarelli 2018: 213, (2)). ShiftP is the projection where the A-Topic is introduced for the first time in the discourse. The G-Topic is introduced below the Ground Phrase (GP) in the Familiar Topic Phrase (FamP). Other projections in (14) are Contrast Phrase (ContrP), Focus Phrase (FocP), and Finiteness Phrase (FinP).

(14) [$_{ForceP}$ [$_{ShiftP}$ [$_{GP}$ [$_{ContrP}$ [$_{FocP}$ [$_{FamP}$ [$_{FinP}$]]]]]]]

Based on the data presented in this section, FamP is a plausible candidate to host an operator that would link the whole clause to the given/familiar content of CG. However, as we have seen, the Topic operator, which is supposed to provide a

---

9 Even though *Ivana* 'Ivan.ACC' is in the clause-final position, it is not Focus. According to Krifka (2007: 18), "Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions." The accusative *Ivana* does not represent an alternative.

referential index to the subject, is not sufficient to license a NS. This restriction applies to both A-Topics and G-Topics:

(15) a.   [$_{ShiftP}$ A-Topic$_i$ [$_{GP}$ [$_{ContrP}$ [$_{FocP}$ [$_{FamP}$ [$_{FinP}$ [$_{TP}$ *pro_i . . .]]]]
     b.   [$_{FamP}$ G-Topic$_i$ [$_{FinP}$ [$_{TP}$ *pro_i . . .]]]]

In fact, this is expected under Frascarelli's Interface Visibility Condition, which states that "minimal overt links optimize the interpretation of Topic chains at the (PF, LF) interfaces" (Frascarelli 2018: 227). According to this condition, the structures in (15) are only possible with overt pronouns in Spec,TP.[10]

The descriptive generalizations in (13) bring us back to Franks' (1995: 307) original insight that referential NSs in Russian (and East Slavic in general) are instances of ellipsis in the context of discursive saliency. In this particular case, we deal with nominal ellipsis, which has recently been brought up again in the literature. Thus, building on Tomioka's (2003) analysis of NSs in Japanese, Barbosa (2019) proposes to link both *pro*-drop and nominal ellipsis to the lack of (article-like) determiners cross-linguistically. If a language lacks (or has a reduced usage of) determiners, several instances of nominal ellipsis are expected to transpire as apparent cases of argument drop, including NS phenomena. Even though East Slavic languages do lack article-like determiners, it does not mean that category D is completely absent in BRU. As it is claimed in this chapter, D is a definitional property of overt pronouns. It should be clarified that, in a realizational model like Distributed Morphology (assumed here), ellipsis is understood as a lack of spell-out (or zero form), rather than post-syntactic deletion at PF. That is, phonological features are assumed to be inserted post-syntactically. From this theoretical perspective, the presence of D is a signal to spell out φP, which would otherwise remain unpronounced, and *pro* is therefore defined as a D-less φP (following Holmberg 2005).

The choice between a DP and a φP is essentially intentional. More precisely, the selection of D for the numeration is not vacuous and is ultimately motivated by the semantic and pragmatic pressures, including the utterer's commitment to the identity of a discursive referent.[11] Thus, a DP is used when the utterer is committed to the identity of the subject, whereas a D-less φP is used if there is no such

---

10 One could argue that an overt pronoun could be base-generated in ShiftP/FamP and linked to *pro* in Spec,TP. However, this option is not plausible because there are no weak or strong pronouns in BRU.

11 Discourse-related decisions at the point of selection can be made in parallel with the syntactic computation, along the lines of Jackendoff's (1997) "parallel architecture" (also assumed in Ackema & Neeleman 2018: 18).

commitment. It seems plausible to suggest that whenever the discursive referent is easily accessible from the context, the utterer does not need to commit to its identity, choosing a φP, instead of a DP. This scenario involves a truncated structure (with just a bare TP or a minimal C-domain), in which φP can be accessed directly from the context without any intervening functional layers:

(16)   CONTEXT: $[_{CP} [_{TP} φP \ldots]$

This structure may have TP-adjoined constituents, but it is incompatible with other fronted material requiring an extended C-domain (see Tsedryk 2015 for further discussion). A truncated structure applies to the clauses in (3), (4), (5b), (10d), and (11b). If this reasoning is on the right track, variability among partial *pro*-drop languages may ultimately fall under an interface theory that is set to model direct accessibility of discursive antecedents, such as Ariel's (1990, 1991) Accessibility Theory (see Gutman 2004 for an implementation).

The structure in (16) is incompatible with Holmberg's (2010a) assumption that φP has to be associated with an A-Topic in the C-domain of the clause (at least for the languages considered in this chapter). I maintain that φP's interpretation can be determined contextually without mediation of the left periphery (but see Madariaga (2022: 75–105) for a different implementation along the lines of Holmberg 2010a and other similar proposals). All in all, there are two options for φP: it is either licensed contextually or bound by a quantifying operator within the clause, receiving a non-referential reading.[12] The second option is another extreme of the utterer's lack of commitment to the identity of the subject and constitutes a lead-in to the discussion of non-referential 3PL subjects.

3 Non-referential 3PL

3PL subjects have a variety of non-referential interpretations, sometimes commonly referred to as "arbitrary" (3PL arbs). For example, Malamud (2013: 4) identifies five such types, using Cabredo-Hofherr's (2003) typology:
(i)   specific existential, (17a);
(ii) vague existential, (17b-c) (including verbs of saying);

---

[12] More precisely, there is also a third option: φP is bound by a local c-commanding antecedent, as in (i).

(i) $[ANTECEDENT_i \ldots [_{CP} [_{TP} φP_i \ldots]]]$

(iii) inferred existential (17d);
(iv) corporate (17e);
(v) universal ("quasi-universal" in Cinque's (1988) terms) (17f).[13]

(17) a. *Pastukali ŭ dzvery. Pajdu adkryju.*
knoked.PL in doors will.go.1SG open.1SG
'Someone has knocked on the door. I will go (and) open.'
(Belarusian)

b. *Motocykl našli vo dvore.*
motorbike.ACC found.PL in courtyard
'The motorbike was found in the courtyard.'
(Russian)

c. *Govorjat', ščo bude došč.*
say.3PL that will.be rain.NOM
'They say it will rain.'
(Ukrainian)

d. *Tut rybu zavaročvali.*
here fish.ACC wrapped.PL
'Here fish was wrapped.' (pointing at a stain
on paper)
(Belarusian)

e. *Sašu otrstranili ot proekta.*
Saša.ACC suspended.PL from project
'They've suspended Sasha from the project.'
(Russian)

f. *U Ispaniji govorjat' po-ispans'ky.*
in Spain speak.3PL in-Spanish
'They speak Spanish in Spain.'
(Ukrainian)

Furthermore, the arbitrary 3PL subject can have a speaker-inclusive reading, but it does not have to. For example, (17f) can be uttered by a Spaniard, but it can also be of any other nationality. There are cases in which the inclusive reading is in fact impossible, which implies that 3PL is not inherently inclusive, consider the following examples (based on Moltmann 2006: 266–267, (23), (26b)).

---

13 The examples in (17) are inspired by Malamud's (2013: 4) examples and discussion.

(18) a. *Kartinu   mogut   uvidet' so   vxoda*.
          picture.ACC can.3PL see.INF from entrance
          'The picture can be seen from the entrance.'

   b. *Ivan   podtverdil, čto kartinu   mogut   uvidet' so*
          Ivan.NOM confirmed   that picture.ACC can.3PL see.INF from
          vxoda.
          entrance
          'Ivan confirmed that the picture can be seen from the entrance.'
          (Russian)

In (18a), the speaker makes an epistemic judgement without belonging to the set of individuals who will potentially see the picture. The same holds for (18b): Ivan is not included in the set, either. It is possible to infer that the speaker's or Ivan's judgement is based on previous experience, but it is just an implicature, which can be cancelled by the follow-up in (19).

(19) *. . .no   ja/on sam eë   ne   uvidel*.
       . . .but   I/he self it   not have.seen
       '. . .but I haven't / he (Ivan) hasn't seen it (from there) myself/himself.'
       (Russian)

In order to have an inclusive experiential reading, the neutral adjectival form has to be used (an equivalent of the expletive 'it is possible'), as shown in (20). Neither of these utterances can be cancelled by (19).

(20) a. *Kartinu   možno   uvidet' so   vxoda*.
          picture.ACC possible.N see.INF from entrance
          'It is possible to / one can see the picture from the entrance.'

   b. *Ivan   podtverdil, čto kartinu   možno   uvidet'*
          Ivan.NOM confirmed   that picture.ACC possible.N see.INF
          so   vxoda.
          from entrance
          'Ivan confirmed that it is possible to/one can see the picture from the entrance.'

As for the overt/null dichotomy, none of the readings in (17), except for the corporate reading, can be obtained with the overt pronoun *jany/oni/vony* 'they' [B/R/U]. If the overt pronoun is used, there is an intended group of people that is referred to. The corporate reading with the overt pronoun is illustrated in (21), where the pronoun is understood as people in the government. This reading comes from

the restriction *va ŭradze* 'in government' in (21a) and the noun *podatky* 'taxes' in (21b). Otherwise, the overt pronoun cannot arbitrarily abstract over people in the same way the 3PL does. As Malamud (2013: 4) points out, "[u]nlike non-arbitrary pronouns, the 3PL arbs in subject position are obligatorily null in Russian."

(21) a. *Jany   va ŭradze   ne   vedajuc', što   robjac'.*
          they.NOM in   government not know.3PL what do.3PL
          'People in the government don't know what they do.'
          (Belarusian)
      b. *Prostym   ljudjam   žyty   nema   na   ščo,   a*
          ordinary.DAT people.DAT live.INF there.is.not on   what and
          vony pidvyščujut' podatky.
          they raise.3PL   taxes.ACC
          'Ordinary people have nothing to live on, and they raise taxes.'
          (Ukrainian)

This dichotomy is further illustrated in (22). In (22a), there are two occurrences of 3PL: crucially, those who say and those who will raise prices form two different arbitrary groups. However, if the overt pronoun is used in the matrix clause, as in (22b), there are two possible readings: (i) those who say and those who will raise prices form the same group, and (ii) those who say and those who will raise taxes are different groups, but in both readings the first group is not an arbitrary one.

(22) a. *Govorjat, čto   budut   podnimat' ceny.*
          say.3PL   that will.be.3PL raise.INF prices.ACC
          'People say that they$_{arb}$ will rise prices.'
      b. *Oni   govorjat, čto   budut   podnimat' ceny.*
          they.NOM say.3PL   that will.be.3PL raise.INF prices.ACC
          'They$_i$ say that they$_i$ will raise prices.'
          'They$_i$ say that they$_{arb}$ will raise prices.'
          (Russian)

More generally, various uses of arbitrary 3PL null subjcts in BRU arise when the utterer does not commit to the subject's identity. The overt 3PL pronoun, on the other hand, arises from the utterer's commitment to the subject's identity. In semantic terms, the first one denotes a property of type ⟨*e,t*⟩, while the second one denotes an entity ⟨*e*⟩. This idea is further developed in the next section.

4 NS at the syntax-semantics interface

4.1 3PL as a function of type $\langle e,t \rangle$

One of the hypotheses that have been proposed in the literature on null arguments is that NS and null objects, in more general terms, instantiate a bare nominal head (Ruda 2017, 2018; Barbosa 2019). In a realizational model of morphology, such as Distributed Morphology, this head is realized as a nominal affix attached to a category neutral root to derive a noun. For example, in words like *rabota* 'work' [Russian/Ukrainian] or *praca* 'work' [Belarusian], this head is realized as *-a* attached to the root *rabot-* or *prac-*.[14] Thus, assuming that such a head (commonly referred to as 'n') can be selected from the lexicon independently from the root, it is proposed that a null argument is the minimal (rootless) nP, as in (23a), where [*u*φ] stands for 'uninterpretable φ-features' (cf. Barbosa 2019: 504, (44)). Furthermore, Barbosa (2019: 503) suggests that (23a) denotes "a property that is trivially true of any individual in the domain", (23b).[15]

(23) a.   nP
        |
       n
       |
     [*u*φ]
    b. $[[nP]] := [\lambda x :\in D_{\langle e \rangle}.x \in D_e]$

Assuming nP, as defined in (23b), I do not take φ-features to be just a set of uninterpretable features checked within the extended projection of nP. I take φ-features to be a function that restricts the domain of application of the function in (23b), yielding a set of discourse participants. For example, the third person would yield a set of individuals that excludes the speaker and the addressee (see Section 4.2 for further details). The plural number indicates that the output set comprises more than one individual. Focusing in this section on 3PL, I take it to be function $\lambda x_e.3PL(x)$, partially building on Barbosa's (2019) analysis of 3PL. More precisely, I maintain that non-referential 3PL in BRU has two ways to be processed at LF, depending on whether the utterer intends to produce an existential (stage-level) statement or wants to achieve a certain level of genericity (with or without spatial unboundedness): 3PL either falls under Existential Closure or ends up in the restrictive clause of the generic operator. These two options are presented in Sections 4.1.1 and 4.1.2, respectively.

---

14 The affix *-a* encodes gender, number and case features.
15 nP semantically corresponds to Elbourne's (2005) null pronoun ONE.

4.1.1 Existential (non-generic) readings

Existential readings are obtained when the 3PL subject semantically incorporates into the verbal predicate and falls under the rule of Existential Closure.[16] Let me elaborate on semantic incorporation in greater detail.

Among possible implementations (see Barbosa 2019: 505 for references), here I follow Farkas & de Swart (2003) in assuming that semantic incorporation is a result of an operation that unifies two functions (predicates), a nominal function and a verbal one. This rule is also known as Unification: "Replace the relevant thematic argument *y* of a verbal predicate with the thematic argument *z* contributed by a nominal argument of the verb." (Farkas & de Swart 2003: 65). These authors use a revised version of Discourse Representation Theory (Kamp & Reyle 1993), abstracting away from the denotation of the branching node that is obtained by Unification. For the sake of explicitness, I restate the definition of Unification in the following terms (using Heim & Kratzer's (1998) style of definition).

(24) *Unification (semantic incorporation)*[17]
       If α is a branching node and {β, γ} is the set of α's daughters, where [[β]] is in $D_{\langle e,t \rangle}$ and [[γ]] is in $D_{\langle e,\langle s,t \rangle \rangle}$, then [[α]] = $\lambda x_e.\lambda e_s.$[[β]](x) $\wedge$ [[γ]](x,e).

This rule makes it possible for a nominal predicate of type $\langle e,t \rangle$ to compose with a verbal predicate of type $\langle e,\langle s,t \rangle \rangle$. The definition in (24) implies that any $\langle e \rangle$-type variable of [[γ]] is replaced by the $\langle e \rangle$-type variable of [[β]] (in line with Farkas & de Swart's (2003) definition). Semantic incorporation is illustrated in (25), where the relevant branching node is VoiceP. As you can see, this node is derived by replacing the argument y_e in Voice′ by the argument x_e of 3PL. Voice′ in its turn is composed by the rule of Event Identification (see the definition in footnote 17).

---

16 *Existential Closure*: For any P ∈ $D_{\langle \sigma,\tau \rangle}$, ∃-closure(P) = $\exists x_\sigma.P(x)$ (based on Heim 1982, cited in Barbosa 2019: 495; here this rule is extended to all types to include events/situations).
17 Unification, as defined in (24), can be compared with Kratzer's (1996) Event Identification:

(i)   *Event Identification*
       If α is a branching node and {β, γ} is the set of α's daughters, where [[β]] is in $D_{\langle e,\langle s,t \rangle \rangle}$ and [[γ]] is in $D_{\langle s,t \rangle}$, then [[α]] = $\lambda x_e.\lambda e_s.$[[β]](x,e) $\wedge$ [[γ]](e).
       (Kratzer 1996: 122; definition taken from Myler 2016: 41)

Event Identification is an operation that unifies two predicates under the identity of two events, whereas Unification is an operation that unifies two predicates under the identity of two individuals. In both cases, the resulting (unified) node is of type $\langle e,\langle s,t \rangle \rangle$.

(25)　　　VoiceP
　　　　$\lambda x_e.\lambda e_s.3\text{PL}(x) \wedge \text{Agent}(x,e) \wedge [[\text{VP}]](e)$　　　Unification

　　3PL　　　　　Voice′
　$\lambda x_e.3\text{PL}(x)$　　$\lambda y_e.\lambda e_s.\text{Agent}(y,e) \wedge [[\text{VP}]](e)$　　Event Identification

　　　Voice　　　　　VP
　$\lambda y_e.\lambda e_s.\text{Agent}(y,e)$　　$\lambda e_s.[[\text{VP}]](e)$
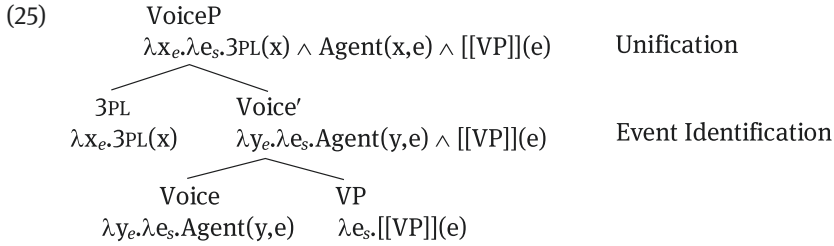
Once 3PL is semantically incorporated into the predicate that is existentially closed at the level of VoiceP, we obtain the LF representation in (26).

(26)　$\exists x_e \exists e_s[3\text{PL}(x) \wedge \text{Agent}(x,e) \wedge [[\text{VP}]](e)]$

Semantic incorporation combined with Existential Closure, applies to specific, vague and inferred existential readings in (17). For example, the sentence in (17b), repeated in (27a), has the interpretation in (27b), which roughly reads: there is an individual x and an event e such that x excludes speech act participants and e is the event of finding the motorcycle by x in courtyard.

(27)　a.　*Motocykl　　　našli　　vo　dvore.*
　　　　　motorbike.ACC　found.PL　in　courtyard
　　　　　'The motorbike was found in the courtyard.'
　　　　　(Russian)
　　　b.　$\exists x_e \exists e_s[3\text{PL}(x) \wedge \text{Agent}(x,e) \wedge \text{found-the-motorcycle}(e) \wedge \text{in-courtyard}(e)]$

Barbosa (2019: 505) points out that one of the consequences of the semantically incorporated 3PL is the semantically neutral number. In fact, the verbal agreement can be plural, but the logical subject does not have to be interpreted as a plurality. For example, (27a) is perfectly fine even if it were just one person who found the motorcycle. The example in (28) further illustrates the number neutrality of 3PL. More specifically, imagine a person calling everyone to the table. Everyone comes, except one person. When she/he finally shows up, (28) is uttered as a reproach for being late. Even though the first verb has plural agreement, this sentence is felicitous if one person called the others to the table.

(28)　*Tebja　　pozvali　k　stolu,　a　　ty　　　ne　idëš'.*
　　　　you.ACC　called.PL　to　table　and　you.NOM　not　come.2SG
　　　　'You were called to the table and you're not coming.'
　　　　(Russian)

4.1.2 Generic (quasi-universal) readings

As Barbosa (2019) claims, individual-level predicates cannot be freely used with the generic 3SG. For example, this is impossible in Finnish and Brazilian Portuguese:[18]

(29) a. * *Brasilia-ssa   rankastaa   samba.*
        Brazil-in   love.3SG   samba
        'In Brazil, one loves samba.'
        (Finnish, Barbosa 2019: 506, (49), citing Anders Holmberg p.c.)
     b. * *Aos   cinquenta anos   sabe   em   quem   confiar.*
        at.the fifty   years know.3SG on   who   trust.INF
        'At age fifty, one knows whom to trust.'
        (Brazilian Portuguese, Barbosa 2019: 507, (50b))

At the same time, Barbosa (2019: 507) observes that individual-level predicates are compatible with arbitrary 3PL null subjects. This is shown in (30) with individual-level predicates such as 'love', 'hate', and 'know'.

(30) a. *U našaj kraine   ščyra   kaxajuc' svajho prèzidènta.*
        in our   country sincerely love.3PL self's   president
        'In our country people sincerely love their president.'
        (Belarusian)
     b. *Na našej kafedre   otkryto nenavidjat dekana fakulteta.*
        on our   department   openly hate.3PL   dean   of.faculty
        'In our department people openly hate the dean of the faculty.'
        (Russian)
     c. *Tut znajut'   tolk   v   pisnjax.*
        here know.3PL appreciation in songs
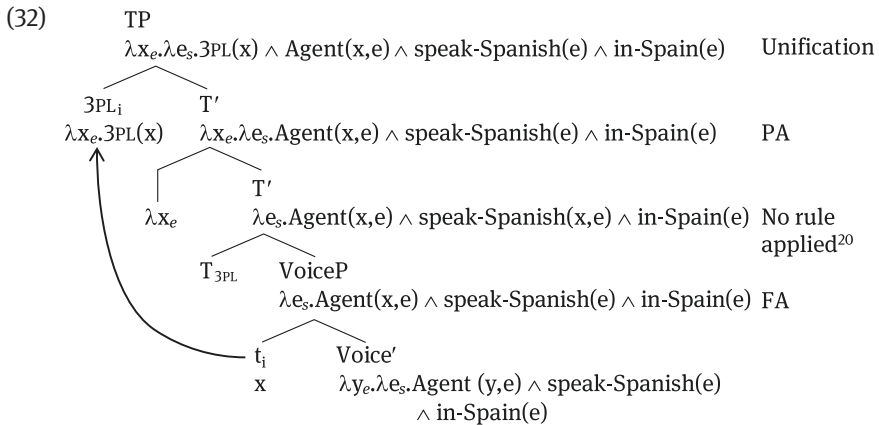   'Here people know how to appreciate songs.'
        (Ukrainian)

In this regard, Barbosa (2019: 508) notes that 3PL may not semantically incorporate into the predicate, in which case the generic operator "quantifies over instances of the kind". I would like to suggest here that semantic incorporation, as defined in (24), still takes place in (30), but it applies in a derived position, namely Spec,TP. Let me illustrate this point on the example in (31), which is the Russian variant of (17f).

---

18 Barbosa (2019: 507) reports that unavailability of the generic 3SG with individual-level predicates is subject to idiolectal variation (it is mostly observed among older speakers).

(31) *V   Ispanii   govorjat'   po-ispanski.*
      in   Spain   speak.3PL   in-Spanish
      'They (people) speak Spanish in Spain.'
      (Russian)

Leaving aside the locative fronting for a moment, we have to focus first on the application of compositional rules at the point of the presumed 3PL movement to Spec,TP. As 3PL is a function, this movement leaves a variable of type $\langle e \rangle$ (i.e., a trace) that is abstracted over by a λ-operator at the landing cite (this is known as Predicate Abstraction; see footnote 19 for a formal definition). An annotated structure is provided in (32). In addition to Unification, applied in the uppermost node, Functional Application (FA) and Predicate Abstraction (PA) are the other two rules in (32) (borrowed from Heim & Kratzer 1998). FA is a basic compositional rule that saturates a predicate by an argument (i.e., the function is applied to an argument of the corresponding semantic type).[19]

(32)

$$\text{TP}$$
$$\lambda x_e.\lambda e_s.\text{3PL}(x) \wedge \text{Agent}(x,e) \wedge \text{speak-Spanish}(e) \wedge \text{in-Spain}(e) \qquad \text{Unification}$$

3PL$_i$ T′
$\lambda x_e.\text{3PL}(x)$ $\lambda x_e.\lambda e_s.\text{Agent}(x,e) \wedge \text{speak-Spanish}(e) \wedge \text{in-Spain}(e)$ PA

T′
λx_e $\lambda e_s.\text{Agent}(x,e) \wedge \text{speak-Spanish}(x,e) \wedge \text{in-Spain}(e)$ No rule applied[20]

T$_{3PL}$ VoiceP
$\lambda e_s.\text{Agent}(x,e) \wedge \text{speak-Spanish}(e) \wedge \text{in-Spain}(e)$ FA

t$_i$ Voice′
x $\lambda y_e.\lambda e_s.\text{Agent}(y,e) \wedge \text{speak-Spanish}(e)$
       $\wedge \text{in-Spain}(e)$

---

19 Definition of the compositional rules in (32) (besides Unification):

(i) *Functional Application*
     If α is the branching node, {β, γ} is the set of α's daughters, and [[β]] is a function whose domain contains [[γ]], then [[α]] = [[β]]([[γ]]).
     (Heim & Kratzer 1998: 44)

(ii) *Predicate Abstraction*
      Let α be a branching node with daughters β and γ, where β dominates only a numerical index i. Then, for any variable assignment a, $[[\alpha]]^a = \lambda x \in D_e . [[\gamma]]^{a^{x/i}}$.
      (Heim & Kratzer 1998: 186)

20 I abstract away from the tense/aspect semantics in (32).

At this point, the following question arises: Is the movement in (32) sufficient to satisfy the Extended Projection Principle (EPP)? If EPP requires a category that can be accessed by the sensori-motor system (i.e., the category that has an overt output at PF), the answer is no, since 3PL is a covert category. In terms of its categorial status, the 3PL null subject is a φP that does not have a D-layer. This layer, as I claim in this chapter, is the category that ensures visibility of φP at PF. In other words, a 3PL null subject cannot satisfy EPP on its own and needs support. Therefore, another XP has to move in order to ensure that the derived structure is properly read at the sensori-motor interface. For concreteness, we could state it in terms of feature checking but nothing hinges on this particular implementation. Suppose that T in BRU has an uninterpretable PHON-feature ('phonological feature').[21] When the 3PL null subject moves to Spec,TP, it cannot check this feature, given that it does not have a D-layer, which supposedly bears an interpretable PHON-feature in BRU. Consequently, another category with an interpretable PHON-feature has to move, which is why the locative is fronted in (31) (see also (30)). This PHON-feature just indicates the potential for sensorimotor accessibility, it does not mean that the actual locative movement has to take place at PF. That is, the locative fronting takes place in narrow syntax and, subsequently, has an impact on LF as well (see also *fn* 21). More precisely, once both the 3PL null subject and the locative adverbial move, they create an input for the generic operator (GN), which independently requires a dyadic structure, composed of a restrictor and a matrix (see Krifka et al. 1995).[22] The corresponding LF is given in (33), where 3PL and the locative expression form a restriction for GN and 'C' stands for a contextually valued variable restricting GN to "appropriate individuals and situations" (Chierchia 1998: 366). The reading is: for any appropriate

---

21 This assumption is reminiscent of Holmberg's (2000) account of stylistic fronting, but there is a crucial difference. Unlike Holmberg, I am not assuming that the PHON-feature of T can only attract the phonological matrix. In fact, in Distributed Morphology (assumed here), phonological features are not part of the narrow-syntactic computation. However, categorial features can "inform" PF about the potential targets of phonological insertion. Thus, the D-layer above φP provides the context for insertion rules operating at PF, or we might say that it has a "phonological signiture" without actually bearing the phonological matrix. From this perspective, the PHON-feature is sensitive to categorial features that have a PF potential (i.e., have an interpretable PHON-feature), but it is not supposed to attarct only the phonological shell.

22 Moltmann (2010: 445) takes the generic operator "to be a combination of a universal quantifier ranging over possible worlds, restricted by some contextually given accessibility relation [. . .], as well as a universal quantifier ranging over individuals." Thus, there are two conditions for this operator to apply: (i) it has to scope over events/situations (or worlds, as maximal situations) and (ii) it requires both a restriction and a nuclear scope. In this regard, it is different from a monadic operator, applied at the level of VoiceP and unselectively binding any variable (see also Mari, Beyssade & Del Prete 2013: 45).

individual x (excluding speech act participants) and event e happening in Spain, it is the case that x speaks Spanish.

(33) GNx$_e$,e$_s$[C(x,e) ∧ 3PL(x) ∧ in-Spain(e)] [Agent(x,e) ∧ speak-Spanish(e)]

The actual semantic contribution of 3PL will be clarified in Section 4.2. Even though it inherently excludes the speech act participants, it can still receive an inclusive reading by pragmatic implicature, provided that the utterer belongs to the reported location (see Section 3).

Now consider the utterance in (34a), which does not have a locative restriction and 3PL has a referential (contextual) reading. Imagine a situation that we are sitting in a bar and hear people talking at the next table. We are trying to understand which language they speak, and then one of us utters either (34a) or (34b).

(34) a. *Govorjat   po-ispanski.*
          speak.3PL   in-Spanish
          'They (people at the next table) speak Spanish.'
      b. *Po-ispanski   govorjat.*
          in-Spanish   speak.3PL
   'They (people at the next table) speak Spanish.'
          (Russian)

In this case, the derivation proceeds as was shown in (32) – with 3PL moving to Spec,TP – but without the locative restriction in the formula. If the PHON-feature of T is not checked by an XP, as in (34a), there is a last resort option to raise V to T after Spell-Out. Normally, the verbal complex [V + Voice] does not raise to T in Russian (Bailyn 2012: 74), but in this particular case, [V + Voice] "hops" to T instead of the affix T-lowering in morphology (at the PF interface). Another option is to have a remnant VP-fronting, as shown in (35a) (the strikethrough indicates an unpronounced copy of V). This second option derives (34b). In any case, the generic operator cannot be applied, because the LF cannot be partitioned into the restriction and the nuclear scope. In (34a), there is no VP-fronting in syntax at all (only V-hopping at the PF-interface), which would preclude GN to have situations in its restriction. In (34b)/(35a), the same ⟨s,t⟩-type predicate (VP) would end up in both the restriction and the matrix, leading to a tautology: for any individual x who speaks Spanish it is the case that x speaks Spanish. Note that Existential Closure cannot be applied either, since 3PL has moved out of VoiceP. The only option is to bind all variables contextually, as presented in (35b), which reads as follows: take individual x (excluding speech act participants) and event e, as supplied by the context, it is the case that x speaks Spanish.

(35)   a.
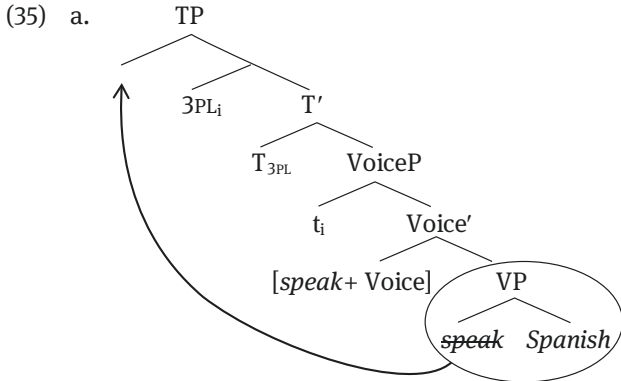
   b.   CONTEXTx$_e$,e$_s$[3PL(x) ∧ Agent(x,e) ∧ speak-Spanish(e)]   cf. (16)

Ultimately, this analysis brings us to the example in (1a), repeated below as (36a). The sentence displays the object-verb word order derived by means of a remnant VP-movement, as shown in (36b). In this case, there is an adverbial within VoiceP, which allows the partition into the restriction (the material moved out of VoiceP) and the nuclear scope (the material contained within VoiceP).

(36)   a.   *Kavu   pjuc'   ranicaj.*
            coffee.ACC   drink.3PL   morning.INS
            'Coffee is drunk in the morning.' / 'One drinks coffee in the morning.'
            (Belarusian)

   b.
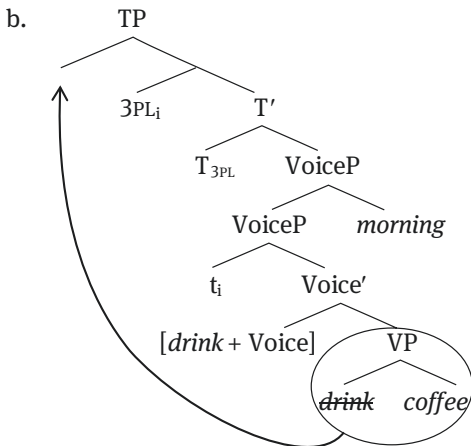
   c.   GNx$_e$,e$_s$ [C(x,e) ∧ 3PL(x) ∧ drink-coffee(e)] [Agent(x,e) ∧ drink-coffee(e)
        ∧ morning(e)]

The LF in (36c) reads as follows: take any appropriate individual x and coffee drinking event e, it is the case that x drinks coffee in the morning.

Before we move on, let me close this section with the following question: Why does the whole VP move in (35b), and not just the direct object or the adverbial? This question brings us to a more general question about the nature of EPP. If we take EPP to be a general [aboutness] feature (i.e., what the sentence is about in Reinhart's (1981) sense) in T, then the answer to the above question lies in the utterer's intention to say something about coffee drinking events, not about a contextually salient entity (i.e. an A-Topic). Frascarelli (2018: 212) takes [aboutness] to be "an extended EPP feature" in the C-domain, but we can put it this way: EPP is a generalized [aboutness] feature in T, while an extended (more precisely, a narrower) instance of [aboutness], reserved specifically for A-Topics (entities), occurs in the extended clausal domain. In other words, the utterances in (30) and (31) are generic statements about a location, whereas the utterance in (36a) is a generic statement about the coffee drinking events, but neither of these utterances has an A-Topic (understood as an entity-type constituent; see footnote 7).

To summarize, 3PL is analyzed as a function of type $\langle e,t \rangle$. It semantically incorporates into the predicate. Semantic incorporation has been defined in (24) as Unification (following Farkas & de Swart 2003). This rule is set to unify a nominal function of type $\langle e,t \rangle$ with a verbal function of type $\langle e,\langle s,t \rangle \rangle$ with the output function of type $\langle e,\langle s,t \rangle \rangle$. 3PL semantically incorporates into the predicate either staying in situ or moving to Spec,TP. The first option leads to an existential reading. The second option leads either to a generic reading (if the clause can be partitioned into the restriction and the matrix) or a contextual reading (if this partition is impossible and the generic operator cannot be applied). The next section further elaborates on the idea that φ-features are functions that restrict the domain of the bare nP.

4.2 φ-features as functions

One of the questions that surrounds the bare-nominal-head approach to NSs concerns agreement in φ-features with the verbal inflection. If the NS is just a nominal head without any features at all, a non-trivial question arises as to the origin of the features in the verbal agreement. As a matter of fact, Ackema & Neeleman (2018: 165) argue that the φ-features can be independently present on the verb and spread to the bare nominal in order to be interpretable. Considering this option in detail would bring us too far afield. If one wants to keep Chomsky's style asymmetric agreement, in which the φ-features of a verb are

valued by those of a nominal, it is inevitable to assume that NSs come in different shapes determined by person and number features. In fact, they are like overt pronouns, except one property: They lack the D-layer in their internal structure (Holmberg 2005). Therefore, it is important to look at the overt pronouns in a given language in order to understand its NSs, if there are any. At the same time, if person and number features can be added to the bare nominal head, yet another question seems pertinent. Namely, how do various feature specifications of a NS affect its denotation? For example, can the first-person singular specification make a NS denote an entity rather than a property? That is, do we have to enrich the inventory of the semantic types as we enrich the inventory of NS φ-features? In this section, I adopt Ackema & Neeleman's (2018) theory of person features, taking a slight detour and bringing the pronominal paradigm of BRU into the spotlight. This theory will be used in the analysis of the generic 2sg in Section 5.1.

Ackema & Neeleman (2018: ch. 2) propose that person features are functions that restrict the universal set of potential referents. This set is composed of two proper subsets which also stand in the set-subset relationship. The smallest subset (S_i) comprises the speaker (i) and her/his context-dependent associates.[23] The larger set (S_{i+u}) includes i, the addressee (u), and their associates. Finally, the largest set of potential referents (S_{i+u+o}) includes i, u, their associates, and the others (o). These sets are organized in an onion-like manner, as visualized in (37).

(37)
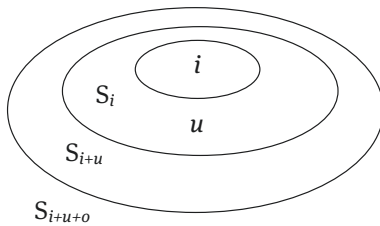
(Ackema & Neeleman 2018: 23, (4))

There are two functions that operate on these sets: the distal function (DIST) and the proximate function (PROX). More precisely, PROX is a function that takes a larger set as its input and delivers the next smaller set as its output. If PROX is applied to S_{i+u+o}, the output would be S_{i+u}. If it is further applied to S_{i+u}, the

---

23 Associates of X are those who are associated with X by some relationship (they can be the same family members, groupmates, team fans, etc.). They are not fixed and vary with the context. See Ackema & Neeleman (2018: 4–8) for further details.

output would be S_i. Basically, PROX removes the complement of a given set. In metaphorical terms, it can be compared with a gradual peeling of an onion: every time it is applied, an outer layer is peeled off, reducing the size of the vegetable. To keep up with the metaphor, the DIST function, on the other hand, operates from the inside of the onion, taking out the core and leaving only the outermost layer. Thus, applied to S_{i+u+o}, it would produce S_o. 'Person', as an umbrella term for these two functions, is essentially an instruction to restrict the input in one way or the other. As for 'Number', it specifies the cardinality value within a given set: either $|S| = 1$ (by default) or $|S| > 1$ (for plural) (here I abstract away from Ackema & Neeleman's proposal to analyze the number in terms of the augmentative feature, see their Chapter 3).

The paradigms of subject person pronouns in BRU are presented in (38).

(38) Subject personal pronouns in BRU

	Belarusian	Russian	Ukrainian
'I'		*ja*	
'you' (2SG)		*ty*	
'he'	*jon*	*on*	*vin*
'she'	*jana*	*ona*	*vona*
'it'	*jano*	*ono*	*vono*
'we'		*my*	
'you' (2PL / polite 2SG)		*vy*	
'they'	*jany*	*oni*	*vony*

The three languages show some variation in the third person, but otherwise the pronominal forms are organized in the same way: the number and person features are fused in one form in the first and the second person; the gender specification (masculine, feminine or neuter) is present in the third person and alternates with the plural specification. The East Slavic languages do not have a special form to differentiate between inclusive and exclusive readings of the first person plural.

In terms of the syntactic structures, the features are organized as shown in (39) (# stands for 'Number' and π for 'Person'). The bare nP (Section 4.1) is taken to denote the largest possible set S_{i+u+o} (N_Π in Ackema & Neeleman 2018: 23).[24]

---

24 See Ackema & Neeleman (2018: 56–57) for motivation behind # being above π.

(39)   Feature structures in BRU

1st person	2nd person	3rd person
#P	#P	#P
# πP	# πP	# πP
| / \	| / \	| / \
(PL) π nP	(PL) π nP	(PL) π nP
|	|	|
PROX	PROX	DIST
|	|	
PROX	DIST	

In (39), I abstract away from Gender (as a node/feature), which could be conflated with # (that is, it should have access to the number specification, so as to allow for the complimentary distribution between the masculine/feminine gender and the plural number in the third person). Furthermore, # could be conflated with π in the first and the second person, but this conflation is supposed to characterize pronominal systems that make the inclusive-exclusive distinction (Ackema & Neeleman 2018: 78–79).

In the structures presented above, the first person is derived by two consecutive applications of PROX, going first from S_{i+u+o} to S_{i+u} (the first application) and then from S_{i+u} to S_i (the second application). The second person is derived by the application of PROX (going from S_{i+u+o} to S_{i+u}), followed by the application of DIST (going from S_{i+u} to S_u). Finally, the third person is derived by the only application of DIST, which allows to shrink the largest set to its outer layer S_o (excluding S_{i+u}). Adding # further specifies whether the number of referents in a given set is greater than or equal to 1. Note that, in this system, what # and π do is restrict a set, but they do not indicate which entity or entities should be picked up from this set. For example, if we take set S_i (derived by π-PROX-PROX) and then add # (set by default to value '1'), the features only instruct that one member should be taken from set S_i, but the features do not instruct which member should be picked up – it could be the speaker or one of its associates. Therefore, an iota-operator is needed to further restrict the set, taking the most salient member in the discourse, namely the speaker. And this is exactly what the D-head is supposed to do by merging with one of the structures in (39). Lumping # and π under the same node φ, we arrive at two pronominal structures, distinguished by their semantic type.

(40) a. Type ⟨e,t⟩

```
        φP
       /  \
      φ    nP
```

b. Type ⟨e⟩

```
           DP
          /  \
         D    φP
             /  \
            φ    nP
```

Furthermore, these structures correspond to the null/overt distinction at the sensorimotor interface, with (40b) being the structure of the overt pronouns in (38). There is nothing new in this dichotomy as long as the overtness is concerned. This is essentially a return to Holmberg's (2005) original conjecture about *pro*, but from a different angle: while distinguishing the null and the overt pronouns by their syntactic weight, this distinction is reframed in the context of the recent proposals with regard to *pro*, as a bare nominal, and φ-features, as restricting functions (Barbosa 2019 and Ackema & Neeleman 2018, respectively).

As was suggested at the end of Section 2, the distinction in (40) is rooted in the utterer's intention to commit or not to commit to the identity of subject. The structure in (40a) is used whenever the utterer is not committed to identify the subject, either relying on the context or not. It has to be licensed by an operator at LF (existential or generic), as discussed in Sections 4.1.1 and 4.1.2. or directly bound by the context (context-forced type-shifting), as suggested in Section 4.1.2 (namely, see (35b)). Section 5 discusses further extensions of this proposal.

5 Extensions

Section 5.1 focuses on 2sɢ, which can have two readings, generic and referential (indexical, addressee-oriented). It will be shown that the generic readings, as in (1b), correlate with the null form, but in many cases, they are also possible with an overt pronoun. The question is why. Section 5.2 discusses inclusive generic readings with *-sja*, as in (1c).

5.1 Generic 2sɢ

In several languages across the world, the second person pronoun is used in generic sentences (Ackema & Neeleman 2018: 123). In this respect, BRU do not constitute an exception. In what follows, I will first describe the distributional properties of the generic 2sɢ (section 5.1.1) and then present its analysis in terms of the person features introduced in section 4.2 (section 5.1.2). Section 5.1.2 is

inspired by Polinsky's (2019) detailed description of the generic 2sg ("arbitrary 2sg" in her terms) in Russian. I will highlight some of the distributional properties that are relevant for the analysis in Section 5.1.2 (the reader is referred to Polinsky's work for a more extensive overview and discussion).

5.1.1 Overtness, speaker inclusiveness, and distributivity

Let us start with a comparison between generic 2sg and the arbitrary 3pl null pronouns. Both are only applicable to human (or human-like), sentient entities, but they also differ in at least three respects. These are (i) alternation with an overt pronoun, (ii) speaker-inclusiveness, and (iii) modal flavours in sentences with negation (Polinsky 2019).

First, the arbitrary 3pl null pronoun is in complementary distribution with an overt form (see Section 3), while its generic 2sg counterpart is not, (41). Note that in (41c), there is also a 2sg generic object in the embedded clause (this 2sg cannot be null).[25]

(41) a. *Žyccë*   (*ty*)   *pačynaeš' calkam razumec'   tol'ki*
          life.ACC   you.NOM   begin.2SG fully understand.INF only
          ŭ   stalyja hady.
          in mature   years
          'You (one) begin(s) to understand the life fully only in mature years.'
          (Belarusian)

      b. (*Ty*)   *snačala bespokoiš'sja, a   potom okazyvetsja vsë*
          you.NOM   first   worry.2SG   but then it.appears   all
          zrja.
          in.vein
          'You worry first, but then it appears that everything is in vein.'
          (Russian)

---

25 In proverbs and sayings, the 2sg generic subject cannot be overt (ia), but this is an idiosyncratic feature of idiomatic constructions. If a non-idiomatic construction of a similar (correlative) form is used, the subject is optionally null (ib).

(i)   a. *Tyxiše   (*ty)   jideš,   dali   (*ty)   budeš.*
          slower   you.NOM go.2SG   further you.NOM will.be.2SG
          'The slower you go, the further you will be.'

      b. *Čym   bil'še (ty)   pracuješ, tym   bil'še xočet'sja   jisty.*
          by.what more   you.NOM work.2SG by.that more   want.3SG.RFL eat.INF
          'The more you work, the more you are hungry.'
          (Ukrainian)

c. (*Ty*)   *nikoly*   *ne*   *znaješ,*   *ščo*   *tebe*   *čekaje*
   you.NOM   never   not   know.2SG   what   you.ACC   await.3SG
   potim.
   afterwards
   'You worry first, but then it appears that everything is in vein.'
   (Ukrainian)

The generic 2SG pronoun cannot be overt if the 2SG specification is the only means to mark genericity in a sentence, which otherwise does not have other formal indications of genericity. In other words, the null form of the generic 2SG optimizes the ambiguity between episodicity and genericity at PF. For example, in (1b), repeated as (42), the use of an overt pronoun would immediately trigger an episodic reading, but the temporal focus in (41a) and (41c) as well as the contrastively coordinated biclausal structure in (41b) are sufficient for signaling a covert generic operator. Thus, proverbs and sayings with a generic 2SG usually have a contrastive biclausal structure (see footnote 25), but the NS is required for further strengthening of their generic force to avoid any possible ambiguity with an episodic reading.

(42)   (**Ty*)   *tut*   *ne*   *porabotaeš'.*
        you.NOM   here   NEG   will.work.2SG
        'It is impossible to work here.' ('One can't (won't be able to) work here.')
        (Russian)

At this point, Polinsky (2019: 407) reports that out of 300 sentences, taken from the Russian National Corpus, 210 have a NS in generic 2SG sentences (overall probability of 0.7).[26] This is a tendency, not a categorical distinction: the 2SG NS increases the chance of the generic reading at the PF interface. The probability of NS is expected to range between 0 and 1 for different types of sentences, depending on their temporal anchoring, contrastiveness and idiomaticity, among other things. In sum, the alternation between the overt and the null form of the generic 2SG is probabilistic in nature, whereas the arbitrary 3PL alternates with a 3PL overt pronoun in a much more deterministic, categorical fashion.

     As for the speaker-inclusiveness, it was shown in Section 3 that the arbitrary 3PL pronoun is inherently speaker-exclusive (inclusive readings can only be prag-

---

26 By the way, the rate of the generic 2SG null object is probabilistically lower than that of the 2SG generic subject (as reported in Polinsky 2019), which, I believe, is due to the agreement-conditioned recoverability (the verb does not show an object agreement in BRU).

matically implicated; see the discussion of (17)–(19)). The generic 2SG pronoun, on the other hand, is always speaker-inclusive.

Finally, with regard to modality in sentences with negation, the sentence in (42) has a modal flavour of circumstantial impossibility/inability, but if the arbitrary 3PL is used, the clause has a flavour of deontic impossibility, consider (43).

(43)   *Tut   ne   rabotajut.*
        here   NEG   work.3PL
        'No one works here.' / 'It is not allowed to work here.'
        (Russian)

As it stands, the generic sentences in (42) and (43) do not have the same compatibility with the available modal bases. Since the generic 2SG pronoun is speaker-inclusive, the modal base in (42) has to be speaker-oriented. That is, the generalization is evaluated relative to the speaker's assessment of abilities and circumstances. In (43), on the other hand, the modal base has to be independent of the speaker: the generalization is evaluated relative to the external system of rules implemented in a given location.

In addition to the three differences between the generic 2SG and the arbitrary 3PL forms mentioned above, some of the key differences between generic 2SG and addressee-oriented 2SG readings should also be mentioned. These include referential vagueness and plurality of possible referents. Thus, unlike the addressee-oriented 2SG interpretation, the generic 2SG interpretation is referentially vague. This property transpires through its inability to bind an addressee-oriented 2SG subject, which, in turn, allows both the generic and the addressee-oriented 2SG readings in the same clause, as in (44a). The addressee-oriented 2SG subject, on the other hand, cannot cooccur with another addressee-oriented pronoun in the same clause, as in (44b).

(44)   a.   *Z   taboj   ne   papracueš.*
           with   you.INS   not   will.work.2SG
           'It is impossible to work with you.' / 'One can't (won't be able to) work with you.'
     b.   **Ty   z   taboj   ne   papracueš.*
           you.NOM   with   you.INS   not   will.work.2SG
           '*You won't be able to work with you.'
           (Belarusian)

Nevertheless, the generic 2sɢ can bind an anaphor, (45a), just like the addressee-oriented 2sɢ, (45b). Note that the generic 2sɢ can be either null or overt in (45a) (compare with (42)).[27]

(45) a. *Sebe* *(ty)* *ne* *obduryš.*
self.ᴀᴄᴄ you.ɴᴏᴍ not will.fool.2sɢ
'You can't fool yourself.'
 b. *Ty* *sebe* *obmanuješ.*
you.ɴᴏᴍ self.ᴀᴄᴄ deceive.2sɢ
'You're deceiving yourself.'
(Ukrainian)

Moreover, despite being morphologically singular, the generic 2sɢ can distribute over a group of referents. This property is illustrated in (46a).

(46) a. *V* *ètom* *tance* *(ty)* *deržiš'sja* *za* *ruki.*
in this dance you.ɴᴏᴍ hold.2sɢ.ʀꜰʟ at hands
'In this dance, dancers hold hands.'
(Polinsky 2019: 398, (22a))
 b. *V* *ètom* *tance* *(ty)* *deržalsja* *za* *ruki.*
in this dance you.ɴᴏᴍ held.ᴍ.ʀꜰʟ at hands
'In this dance, you were holding (someone's) hands.' (no distributive reading)
(Russian)

In (46a), to be read as an instruction for a dance, the event of holding hands is distributed over as many dancers as required by the dance (with two dancers the reading would be reciprocal). In (46b), which is an episodic utterance featuring the addressee-oriented 2sɢ subject, the distributive reading is impossible.
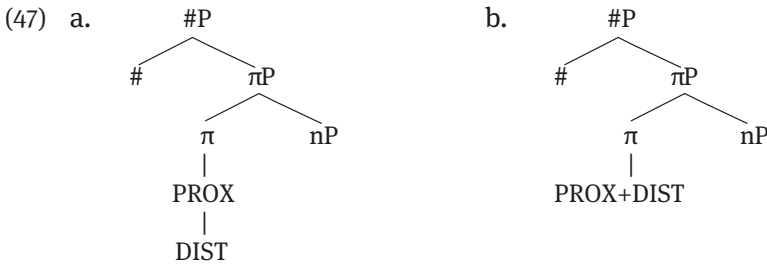
In conclusion, the generic 2sɢ subject can be either overt or null (depending on the probabilistic ambiguity with the episodic reading). Unlike the 3ᴘʟ arbitrary subject, it is inherently speaker-inclusive, and, unlike the addressee-oriented 2sɢ, it entails a plurality of referents in a given domain despite its morphologically singular number.

---

27 I think the contrast between (42) and (45a) (with respect to the overtness of the generic 2sɢ subject) is due to the spatial boundedness. The sentence in (42) is spatially bounded, which decreases its level of genericity. Therefore, the NS is required for the optimization of the generic reading at PF. The sentence in (45a), on the other hand, is spatially unbounded, which weakens the chance of ambiguity with an episodic reading with an overt pronoun.

5.1.2 Unordered PROX and DIST

Returning now to the system of features/functions presented in Section 4.2, I follow Ackema & Neeleman (2018: 125), who propose that generic 2SG subjects have the same person features as regular 2SG subjects do, namely PROX and DIST, but these features are not ordered with respect to each other. In other words, there are two possible ways in which these features can be organized, as shown in (47) ((47a) repeats the 2nd person structure from (39)).

(47)   a.

```
            #P
          /    \
        #       πP
               /   \
              π      nP
              |
            PROX
              |
            DIST
```

b.

```
            #P
          /    \
        #       πP
               /   \
              π      nP
              |
          PROX+DIST
```

An ordered application of PROX and DIST, as in (47a), ends up in set S_u (including the addressee and the addressee's associates, but excluding the speaker, the speaker's associates, and the others). Recall that the application of PROX peels off the outer layer of the initial set S_{i+u+o}, leaving S_{i+u}. The remaining set is the input for the subsequent application of DIST, which takes out the core and leaves the outermost layer, S_u. In (47b), PROX and DIST each apply to the initial set S_{i+u+o}, resulting in a unification of two output sets. This application proceeds in the same way as function addition in mathematics, as formalized in (48), where F stands for 'feature/function'.

(48)   If $F_1(S) = S_1$ and $F_2(S) = S_2$, then $F_1+F_2 = S_1 \cup S_2$.
        (Ackema & Neeleman 2018: 125, (40))

When PROX is applied to S_{i+u+o}, it delivers set S_{i+u}. Also applied to S_{i+u+o}, DIST delivers set S_o (leaving only the outer layer). According to (48), the result of PROX+DIST in (47b) is $S_{i+u} \cup S_o$, which is the initial set S_{i+u+o}. This is the most inclusive set of person referents, comprising the speaker, the addressee, their associates and the others, which characterizes the generic nature of the structure in (47b). This is how Ackema & Neeleman's (2018: 123–127) proposal accounts for the speaker-inclusiveness described in Section 5.1.1.

Furthermore, the largest set delivered by the application of [PROX+DIST] comprises a plurality of possible referents. The singular number (encoded by

#) only instructs that one member of this set can be selected for interpretation, but it does not specify which one. The generic operator distributes over any possible member of the set, resulting in a sum of singular entities, which accounts for the distributive reading in (46a). As for the addressee-oriented pronoun *ty* 'you' in BRU, it has the structure in (47a) merged with the D-head. Recall that the function of D is to pick up the most salient referent out of set S_u (which includes the addressee and his/her associates). That is, the overt pronoun *ty* is a DP that has the addressee as its only referent, which is why the addressee-oriented *ty* is incompatible with the distributive readings, as was shown in (46b).

As suggested in Section 4.2, the syntactic distinction between a D-less structure of φ-features (φP) and a DP has a direct effect at the interfaces: it corresponds to the null/overt distinction at PF and to the property/entity distinction at LF. That is, each overt pronoun of type $\langle e \rangle$ has a corresponding null pronoun of type $\langle e,t \rangle$. This dichotomy lies at the heart of the structures in (39), where we have three arrangements of person features: PROX–PROX, PROX–DIST and DIST, corresponding to sets S_i, S_u and S_o, respectively. This is where we have a one-to-one correspondence between overt and null forms in parallel with D and a D-less φP.

Now, the fourth arrangement, PROX+DIST in (47b), has just been added. It corresponds to the least restricted set S_{i+u+o}, in which neither the speaker, the addressee, nor anybody else is supposed to be singled out. That is, the D-head is functionally useless in this case and the structure (47b) automatically falls outside the D-less/D, property/entity or null/overt dichotomy. In other words, the φ-structure in (47b) can be either overt or null without D, and we thus derive that the overt and the null instances of the generic 2SG are not distinguished categorially (their optionality is constrained by a variety of interface factors, as was mentioned in Section 5.1.1). The arbitrary 3PL, on the other hand, is a D-less φP, being opposed to the same φP that is merged with D.

To conclude, there are three types of spell-out rules for the subject pronouns in BRU: two categorially conditioned rules in (49a) and (49b) (D/D-less), covering the paradigm in (38) and the φ-feature structures in (39), and a categorially undefined rule in (49c), specific to (47b), with an optional spell-out. In any case, the overt 2SG form (*ty* 'you') spells out PROX and DIST features (whether ordered or not) in the context of # (without PL).

(49) a. [$_{DP}$ D [$_{φP}$ φ [$_{nP}$ n]]] $\longleftrightarrow$ one of the forms in (38)
      b. [$_{φP}$ φ [$_{nP}$ n]] $\longleftrightarrow$ $\varnothing$
      c. [$_{φP}$ φ [$_{nP}$ n]] $\longleftrightarrow$ $\varnothing$ or overt 2SG, if π is PROX+DIST

In Section 5.2 I consider yet another possibility (there are no φ-features at all, only a bare nP), finally brining up the last item of the paradigm in (1), repeated below in (50).

5.2 Genericity with *-sja*

It is far beyond the scope of this chapter to provide an extensive discussion of constructions like (50) in BRU. My goal is solely to show how (50) (and other such constructions) could fit the whole picture provided thus far.

(50)   *Tut   pracjujet'sja   dobre.*
      here   work.3SG.RFL   well
      'It feels good to work here.' / 'One feels good working here.'
      (Ukrainian)

In (50), the adverb *dobre* 'well' does not describe the manner of the eventuality, but a psychological disposition towards the eventuality.[28] That is, it is a disposition adverb, rather than a manner adverb (see Kim & Siloni 2020: 274–275). An overt dative subject is possible in (50), but it would turn down the person-related genericity. Furthermore, the presence of an overt dative makes the locative optional (51a). However, the locative expression cannot be dropped if the person-related genericity is to be preserved (51c).[29] The disposition adverb cannot be dropped even when an overt dative is present (51b). Also note that the locative expression can form a clause with the disposition adverb independently from the verb with *-sja*, (51d). Again, the locative expression cannot be dropped under the intended generic reading, (51e). Finally, an overt dative argument is possible without the verb, (51f), as long as the manner adverb is present, (51g).

(51) a.   *Meni   (tut) pracjujet'sja   dobre.*
         me.DAT here   work.3SG.RFL well
         'I work well (here).' / 'I feel good working (here).'
   b.   *Meni   pracjujet'sja   *(dobre).*
         me.DAT work.3SG.RFL well
         'I work well (here).' / 'I feel good working (here).'

---

28 Constructions like (50) are known as "dispositional constructions" (Kim & Siloni 2020).
29 Alexiadou & Carvalho (2017) argue that the locative PP is the external argument in generic *se*-constructions in Brazilian Portuguese.

c. *Pracjujet'sja dobre.*
   work.3SG.RFL   well
   'One works well.'

d. *Tut   duže   dobre.*
   here   very   well
   'One feels very good here.'

e. *Duže dobre.*
   very   well
   'One feels very good.' (generic)

f. *Meni   (tut)   duže   dobre.*
   me.DAT   here   very   well
   'I feel very well here.'

g. *Meni   *(duže dobre).*
   me.DAT   very well
   'I feel very well'
   (Ukrainian)

The distribution presented above shows at least two things about (50): (i) the dative argument has nothing to do with the verb (but it is associated with the disposition adverb, see also Rivero & Arregui 2012: 314), and (ii) there are two types of predication: the verbal one and the one that independently involves the locative expression and the disposition adverb; the verb does not project the same structure as it would without *-sja*.

According to Kim & Siloni (2020: 263), who adopt Reinhart's (2016) Theta System, the common denominator of the verbs that constitute the input for the dispositional constructions like (50) is the presence of a positive cluster of atomic theta-features [+m] and, optionally, [+c]. As Kim and Siloni (2020: 259) put it, "[t]he feature [c] determines whether or not the argument in question is necessarily responsible for causing the denoted event (or change). The feature [m] determines whether or not the mental state of the argument in question is relevant for the denoted event, that is whether the event involves volition and intention of the argument.". More precisely, they suggest that the input verb, specified as [+c+m] or [+m], undergoes a feature adjustment and ends up with the cluster [−c+m]. Because of these features, the verb does not have VoiceP in its extended projection. Therefore, I suggest that the [+m] feature is minimally realized by the nP spelled out as *-sja*.[30]

---

30 Kim & Siloni (2020) argue that the feature adjustment from [+c+m]/[+m] to [−c+m] is a strictly lexical operation, which has nothing to do with syntax. This idea, however, is not necessarily incompatible with late-insertion in the spirit of Distributed Morphology. Thus, we could still

(52)
```
              VP
            /    \
         nP        V
         |       [−c+m]
        sja
```
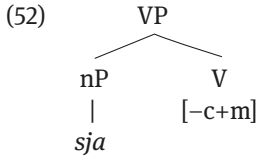
Recall that nP denotes a property (a set of entities) – it is of type $\langle e,t \rangle$ – and it does not fully saturate the predicate (it combines by Unification, defined in (24), as opposed to a DP argument, of type $\langle e \rangle$, that would be merged by Functional Application in Spec,VoiceP, if the verb had a full-fledged projection). In other words, I propose that the merger of nP (i.e., syntactic realization of a thematic feature) is a consequence of the lexical feature adjustment, but the morpheme *-sja* itself does not trigger this adjustment.

As assumed in Section 4.2, nP (realized as *-sja* in this particular case) denotes the largest possible set of person referents, S_{i+u+o}, making the VP in (52) compatible with the generic quantification. At this junction, two options should be considered, i.e. (i) either the predicate is saturated by a dative argument (after the VP is merged with a disposition adverb), or (ii) the nP falls under the scope of the generic operator; this operator independently requires a restrictive clause, which is expressed by the locative expression in (50). The locative expression, therefore, cannot be dropped in the absence of the dative argument, as shown in (51c); otherwise, the locative expression is optional when the dative argument is merged, as in (51a). Finally, the verb takes the 3sg agreement by default.

6 Conclusion

The goal of this chapter was to provide a unified analysis of various instances of East Slavic null subjects in root clauses. Following Holmberg (2005), the null/overt dichotomy has been maintained along the lines of the syntactic weight of a pronoun, namely φP versus DP (assumed in the literature since Cardinaletti & Starke 1999 and Déchaine & Wiltschko 2002). This idea has been implemented in the context of the current work on null subjects (Frascarelli 2018; Barbosa 2019) and person features (Ackema & Neeleman 2018). A pronoun is assumed to be composed of a bare nominal head (which does not categorize any lexical root), subsequently equipped with person and number features.

---

assume that the verb is projected from a root, which is categorized by 'little v'. In this approach, v could simply have different combinations of [m] and [c] features (taken from the lexicon, not syntactically derived). In (52), for example, it would be v[−c+m] merged with a category-neutral root.

Person features are functions that are applied to a set of possible person referents (Ackema & Neeleman 2018). The bare nominal head, syntactically a bare rootless nP, denotes the largest possible (the most inclusive set), which is further restricted by such person features, as 'proximate' (PROX) and 'distal' (DIST). Languages vary along the lines of possible ways to apply PROX and DIST. In East Slavic, theses features/functions are applied in four ways: (i) ordered double application of PROX (first person), (ii) ordered application of PROX and DIST (second person), (iii) unordered application of PROX and DIST (generic 2nd person), and (iv) a single application of DIST (third person). East Slavic languages do not have a single application of PROX, which would derive a generic inclusive pronoun 'one'. This is a parametric choice of these languages, which is functionally compensated by the four available ways to apply PROX and DIST. In this respect, East Slavic languages functionally resemble the consistent NSLs in their ability to compensate the lack of 'one' by using the arbitrary 3PL, the generic 2SG and *-sja*, as was shown in (1). However, this is just a surface resemblance that does not bear on their status of being partial NSLs.

Partial *pro*-drop is, among other things, an interface phenomenon that is driven by the visibility condition imposed on minimal topic links within a clausal domain (Frascarelli 2018). It has been shown that this condition applies to both A- and G-Topics and, furthermore, that there are two options for *pro*, specifically (i) it stays in Spec,VoiceP and falls under Existential Closure, or (ii) it moves to Spec,TP, in which case it either receives a generic reading, being part of the restriction of the generic operator, or is contextually bound. EPP (as a property of T) has been defined in terms of a general [aboutness] feature which is not restricted to entities. An entity-specific occurrence of [aboutness], reserved for A-Topics, is found in the extended C-domain (Frascarelli 2018). In addition, T in East Slavic requires a category that can be accessed by the sensorimotor system. D has been claimed to be such a category. As far as the PF interface is concerned, the East Slavic languages show a consistent correlation between the presence of D in syntax and the overt form (vocabulary insertion). In other words, spell-out is categorially conditioned, except for the generic 2SG, which is the only instance of the null/overt alternation associated with a D-less pronominal structure. Based on the discussion presented in this chapter, there are four spell-out rules in BRU, as summarized in (53) and (54), partially building upon (49).

(53) a. $[_{DP}$ D $[_{\varphi P}$ φ $[_{nP}$ n$]]]$ ⟷ overt pronoun
    b. $[_{\varphi P}$ φ $[_{nP}$ n$]]$ ⟷ ∅

(54) a. $[_{\varphi P}$ φ $[_{nP}$ n$]]$ ⟷ ∅ or overt 2SG, if π is PROX+DIST
    b. $[_{nP}$ n$]$ ⟷ *-sj*a / ___V[−c+m]

The rules in (53) are general rules replicating Holmberg's (2005) overt/null dichotomy. The rules in (54) are specific, i.e. (54a) is conditioned by a special arrangement of person features and (54b) is lexically conditioned rule and applied to a bare nP in the context of a dispositional verb, as discussed in Section 5.2. Whether -*sja* spells out nP across the board in East Slavic languages remains an open question.

All in all, partial NS phenomena represent a complex interaction of various factors operating at the interface levels, including the universal inventory of person features, their specific arrangements and spell-out rules, which interact with some (more general) interface visibility conditions on the [aboutness] feature at the level of TP or in its extended C-domain.

References

Ackema, Peter & Ad Neeleman. 2018. *Features of person: From the inventory of persons to their morphological realization*. Cambridge, MA: MIT Press.

Alexiadou, Artemis & Janayna Carvalho. 2017. The role of locatives in (partial) *pro*-drop languages. In Michelle Sheehan & Laura R. Bailey (eds.), *Order and structure in syntax II: Subjecthood and argument structure*, 41–67. Berlin: Language Science Press.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, Mira. 1991. The function of accessibility in a theory of grammar. *Journal of Pragmatics* 16. 443–463.

Bailyn, John. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press.

Barbosa, Pilar. 2019. *Pro* as a minimal nP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50 (3). 487–526.

Bianchi, Valentina & Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia* 2. 43–88.

Bizzarri, Camilla. 2015. Russian as a partial *pro*-drop language. Data and analysis from a new study. *Annali di Ca'Foscari* 49. 335–368.

Cabredo-Hofherr, Patricia 2003. Arbitrary readings of third person plural pronominals. In Matthias Weisgerber (ed.), *Proceedings of the conference "sub7-Sinn und Bedeutung"*, 81–94. University of Konstanz.

Cardinaletti Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk Van Riemsdijk (ed.), *Clitics in the Languages of Europe*, 145–233. Berlin: Mouton de Gruyter.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6. 339–405.

Cinque. Guglielmo 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19. 521–581.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33. 409–442.

Elbourne, Paul. 2005. *Situations and individuals*. Cambridge, MA: MIT Press.

Farkas, Donka & Henriëtte de Swart. 2003. *The semantics of incorporation*. Stanford, CA: CSLI Publications.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.

Frascarelli, Mara 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistics Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null-subject languages. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: A synchronic and diachronic perspective*, 211–239. Oxford: Oxford University Press.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: John Benjamins.

Gruber, Bettina. 2017. Temporal and atemporal uses of 'you': Indexical and generic second person pronouns in English, German, and Dutch. *Journal of Comparative Germanic Linguistics* 20. 199–227.

Gutman, Einat. 2004. Third person null subjects in Hebrew, Finnish and Rumanian: An accessibility-theoretic account. *Journal of Linguistics* 40, 463–490.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite NPs*. Amherst, MA: University of Massachusetts PhD dissertation.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31 (3). 445–483.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36 (4). 533–564.

Holmberg, Anders. 2010a. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. The null generic subject pronoun in Finnish: A case of incorporation in T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders & Michelle Sheehan. 2010. Control into finite clauses in partial null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–152. Cambridge: Cambridge University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Potuguese, Finnish and Marathi. *Studia Linguistica* 63 (1). 59–97.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kamp, Hans & Uwe Reyle. 1993. *From Discourse to Logic. Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Kim, Dganit Jenia & Tal Siloni. 2020. The dative dispositional construction. *Linguistic Inquiry* 51 (2). 237–279.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase Structure and the Lexicon*. Dordrecht: Kluwer.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Working Papers of SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 6*, 13–56. Potsdam: Universitätsverlag Potsdam. https://www.sfb632.uni-potsdam.de/images/isis/isis06.pdf (accessed May 23, 2020).

Krifka, Manfred, Francis J. Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: An introduction. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The Generic Book*, 1–124. Chicago: University of Chicago Press.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.

Madariaga, Nerea. 2018. Diachronic change and the nature of pronominal null subjects: The case of Russian. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: Synchronic and Diachronic Perspectives*, 171–198. Oxford: Oxford University Press.

Madariaga, Nerea. 2022. Referential null subjects in Russian. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.), *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure, and Typology*, 75–104. Berlin: Mouton De Gruyter.

Malamud, Sophia. 2013. (In)definiteness-driven typology of arbitrary items. *Lingua* 126. 1–31.

Mari, Alda, Claire Beyssade & Fabio Del Prete. 2013. Introduction. In Alda Mari, Claire Beyssade & Fabio Del Prete (eds.), *Genericity*, 1–92. Oxford: Oxford University Press.

Moltmann, Friederike 2006. Generic one, arbitrary PRO and the first person. *Natural Language Semantics* 14. 257–281.

Moltmann, Friederike. 2010. Generalizing detached self-reference and the semantics of generic *one*. *Mind and Language* 25(4). 440–473.

Myler, Neil. 2016. *Building and Interpreting Possessive Sentences*. Cambridge, MA: MIT Press.

Polinsky, Maria. 2019. When agreement and binding go their separate ways: Generic second person pronoun in Russian. In Matthew Baerman, Oliver Bond & Andrew Hippisley (eds.), *Morphological perspectives: Papers in Honour of Greville G. Corbett*, 390–413. Edinburgh: Edinburgh University Press.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27: 53–94.

Reinhart, Tanya. 2016. The Theta System: Syntactic realization of verbal concepts. In Martin Everaert, Marijana Marelj & Eric Reuland (eds.), *Concepts, Syntax, and Their Interface: The Theta System*, 1–111. Cambridge, MA: MIT Press.

Rivero, María Luisa & Ana Arregui. 2012. Building involuntary states in Slavic. In Violeta Demonte & Louise McNally (eds.), *Telicity, change, and state: A Cross-categorial View of Event Structure*, 300–332. Oxford: Oxford University Press.

Ruda, Marta. 2017. *On the Syntax of Missing Objects: A Study with Special Reference to English, Polish, and Hungarian*. Amsterdam: John Benjamins.

Ruda, Marta. 2018. Syntactic representation of null arguments. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: Synchronic and Diachronic Perspectives*, 240–260. Oxford: Oxford University Press.

Sgall, Peter, Eva Hajičová & Jarmila Panenová. 1986. *The Meaning of the Sentence and Its Semantic and Pragmatic Aspects*. Dordrecht: Reidel.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In Kerstin Schwabe & Susan Winkler (eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*, 321–340. Amsterdam: John Benjamins.

Tsedryk, Egor. 2015. Deriving null pronouns: A unified analysis of subject drop in Russian. In Malgorzata Szajbel-Keck, Roslyn Burns & Darya Kavitskaya (eds.), *Formal Approaches to Slavic Linguistics 23: The First Berkeley Meeting*, 342–361. Ann Arbor: Michigan Slavica Publication.

Nerea Madariaga

3 Referential null subjects in Russian: A synchronic and diachronic overview

1 Introductory remarks

In this chapter, I will offer an overview of null subjects (NSs) with individual reference interpretation in the history of Russian.[1] Specifically, I will relate the current use of NSs in root contexts and NSs in syntactically embedded/subordinate clauses in Russian to their diachronic development.

The diachronic approach of this chapter will take as its starting point the synchronic description of the two relevant linguistic systems representing the initial and final stages of the change, i.e. referential NSs in Old Russian and Modern Russian. Fortunately, detailed synchronic accounts of these elements have been previously put forward, especially for Modern Russian. The main contribution of this chapter will then be to relate the two NS systems, which are, in principle, very different, and explain the reasons for a natural transition from one grammar to the next.

To be more precise, I will take every silent subject to be a minimal φP, in the spirit of Holmberg (2005, 2010), endowed with unvalued φ-features that, in addition, needs to copy the referential index of some D-valued antecedent, in order to be interpreted as referential (otherwise, it would be interpreted as generic or arbitrary).

On the one hand, I will argue that Old Russian was a "well-behaved" *pro*-drop language (Meyer 2011; Madariaga 2015, 2018; Jung 2018). Old Russian will represent our Grammar 1, in which NSs were licensed in an "indirect" way, in the domain of a T, whose unvalued D-feature was valued by a null Topic in CP (Frascarelli 2007; Holmberg, Nayudu & Sheehan 2009), whereas emphatic or contrastive subjects had to be realized as overt pronouns.

Pronominal NSs were licensed in a free and uniform way, regardless of the syntactic context (root or embedded), and their specific topical status (familiar/ given or sentential). The relevant configuration for licensing NSs was underpinned by the existence of V-to-T movement in Old Russian, which ensured the ability of T to check the EPP, match φ-features *and* value its unvalued D-feature from that position, thus becoming suitable to transmit that D-feature to the NS as well. As expected in a consistent NSL, other arguments in the structure lacking this Agree relation with T could not be dropped in Old Russian, but had to be overtly realized as clitics.

Our Grammar 2 will be the next stage in the change, starting from Late Middle Russian and represented by Modern Russian. This language has been characterized as a partial NS language, e.g. by Holmberg, Nayudu & Sheehan (2009); Roberts & Holmberg (2010); Livitz (2014); Tsedryk (2012, 2015, 2022), and many others, following the classification by Holmberg (2005). The baseline realization of pronominal subjects is overt in partial NSLs, although they can be dropped under certain circumstances, which vary from language to language. Here, I take "partial" NSL to be a cover term for languages displaying a range of properties that distinguish them from consistent NSLs.

For the purposes of analyzing NSs with individual reference interpretation, a significant property will be the availability of null Topics in these languages, which includes all sorts of null arguments, although we will just focus on null topic subjects in this chapter. The relevant configuration to license referential NSs in Modern Russian is the lack of V-to-T movement (Bailyn 2012; Gribanova 2013), together with a T-head lacking an unvalued D-feature, which is by itself able to license only non-referential NSs (Holmberg, Nayudu & Sheehan 2009). The definite (referential) interpretation of a NS requires an independent mechanism that can be obtained from heterogeneous sources. All of them imply the existence of an additional grammatical relation that is established with some higher element through CP. This can be performed by transmitting a D-feature from a preceding sentential Topic (A-Topic), from an Edge-feature (a logophoric feature or a situational/contextual topic feature, which renders a familiar/given topic, i.e. G-Topic), or from a c-commanding antecedent through a bound variable (embedded "control"). Impressionistically speaking, NSs in Modern Russian are less homogeneous, albeit more restricted, than NSs in Old Russian.

In a nutshell, I will argue that Russian first lost V-to-T movement; this loss conveyed the inability of T to mediate in copying the D-feature that referential NSs needed from the element located in CP. I formalize this change as a reanalysis of T losing its unvalued D-feature. This new configuration forced pronominal referential subjects to be realized overtly in regular conditions, which means that Russian factually became a non-*pro*-drop language. However, NSs with individual

reference were still available to new learners in their linguistic input because (i) change spreads through a community of speakers in a progressive way, and (ii) infinitive clauses still produced "referential" gaps in the place of subjects (PRO) in certain contexts.

I will propose that the instances of NSs we find in Modern Russian emerged by reanalysis/extension of residual NSs inherited from the previous system, and, in particular, NSs in root contexts were reinterpreted as null Topics (G-Topics or successive instances of an A-Topic), whenever their reference could be linked directly to the right topic or topic feature at CP; and NSs in embedded contexts were subsumed under "control" requirements (the need of a local c-commanding antecedent), extending the rare Old Russian instances of control which had been previously available only in a few infinitive constructions.

In this chapter, I will incorporate some basic assumptions about the nature of silent pronouns and the relation of T with NSs. These are listed below.

First of all, I will follow Holmberg (2005) and related work, such as Holmberg, Nayudu & Sheehan (2009), Roberts (2010); Livitz (2014) and Tsedryk (2015), by assuming that NSs in Russian are minimal φPs with unvalued interpretable φ-features ($φP[_{iφ\_}]$). When they have their φ-features matched at [Spec,TP], they can exhibit gender/number verbal agreement, but they still need to match an additional D-feature to be interpreted as referential.[2]

Second, following Holmberg, Nayudu & Sheehan (2009), I will assume that T comes in two flavors. Specifically, in consistent NS languages, it contains an unvalued D-feature, which can be satisfied by an overt (in)definite DP or a NS (φP) through a null Topic in CP, with the ability of endowing the NS with a definite interpretation.[3] In partial NS languages, however, T lacks a D-feature, so NSs, after valuing their φ-features in [Spec,TP], are canonically interpreted as generic or arbitrary, unless they are able to copy the referential index of a valued DP, by entering either in a chain with it (a "topic chain"), or an anaphoric relation (getting bound by a c-commanding antecedent).[4] As a third option available for

---

2 I diverge at this point from Tsedryk (2022), who adopts Holmberg's (2005) idea that referential NSs are full DPs, combined with Barbosa's (2019) account based on semantic types.

3 For the purposes of this chapter, I do not need to assume a rich extended projection of CP in Old Russian (our consistent NS language), but just assume that NSs are uniformly linked to any type of null operator / topic in the C-domain (regardless of the degree of embedding), just as far as they share a [+given] feature with it (see Jiménez-Fernández 2016 for Spanish). I will explain this in more detail in Section 3.

4 The presence of a D-feature on T implies that the Avoid Pronoun Principle proposed by Chomsky (1981) is active only in the case of consistent NS languages. In a context in which either a null or overt pronoun is available, the null variant is preferable for expressing referential dependence

NSs, I will follow Sigurðsson (2011) and Tsedryk (2015) in that specific informational features (Edge-features), such as logophoric and situational/contextual/topical (deictic) features, can also transmit a referential index to a D-lacking φP, rendering it definite/referential.

With this line of thought in mind, in section 2 I review the synchronic properties of referential NSs in Russian, both in root clauses (section 2.1) and embedded clauses (section 2.2). In section 3, I deal with the properties of NSs in Old Russian. Then, in section 4, I explain the diachronic process that determined the current shape of the Russian system of referential NSs nowadays (the transit from Grammar 1 into Grammar 2). Finally, section 5 offers a conclusion.

2 Modern Russian NSs with individual reference (Grammar 2)

2.1 Root contexts and null Topics

In this section, I describe the use of NSs in root contexts in Modern Russian and introduce some basic assumptions. I argue that, in Modern Russian, NSs with the individual reference interpretation in root sentences have two possible sources, both tied to the C-domain: (i) null G-Topics, also known as Familiar Topics; and (ii) successive occurrences of an A-Topic (of a Sentence Topic). Every other instantiation of a subject with individual reference must be overtly realized in root contexts.

Russian shares with other partial NS languages their most salient common property, which is usually defined by opposition both to "*pro*-drop" and "non-*pro*-drop" languages (Holmberg 2005), i.e. the baseline realization of pronominal subjects is overt, *but* there are several contexts in which subjects can also be null (there is *pro*-drop to some extent).

One of the conditions determining licensing of NSs in Russian, as well as in other partial NS languages, is the availability of null Topics (of the type analyzed by Sigurðsson 2011 in Germanic). For the purposes of this chapter, we need to distinguish at least two main types of topic subjects in Russian, which can be described in terms of recoverability of the reference of the subject and

---

only if T has a D-feature. Otherwise, by default, a null variant will express genericity or have an arbitrary interpretation.

its interpretation.[5] While in some cases the reference of a NS can be recovered from the "context" (see (1a–b)), in other cases there seem to be some additional restrictions. In the latter configurations, the reference of a silent subject cannot be recovered even from the preceding context, and an overt pronoun must surface instead (1c).

(1) a. *Privet! Xorošo, čto (vy)   prišli!*   (Logophoric use)
         hi   well   that you.NOM came.PL
         'Hi! So good that you came!'
    b. A: *Čto   s   otcom?*   (G-Topic)
           what with father
           'What's up with father?' / 'Anything new about father?'
        B: *(On) vsë eščë boleet.*
           (he) all still is_sick
           'He's still sick.' (Tsedryk, 2022: 35–75, example (5))
    c. A: *Ty čto-nibud' znaeš' ob Ivane?*   (A-Topic)
           'Do you know anything about Ivan?'
        B: *(On)   sejčas rabotaet   na zavode.
           he.NOM   now   work.3SG   on factory
           'He is now working in a factory.' (adapted from Tsedryk, 2022: 35–75, example (8))[6]

The subjects in (1a) and (1b) are interpreted as Krifka's (2007) familiar Topics (also known as G-Topics), which are part of the discourse common ground, shared by the speaker and the hearer. They are semantically present in the context, either in utterances pronounced in the preceding discourse or just in previous illocutionary acts (1b). I also include topics related to logophoric features in this group (1a).[7] If we take a look at example (1b), the question formulated by the first

---

5 I will not discuss the role of the register in this chapter but we can say that, in general terms, colloquial Russian is more prone to drop pronominal subjects than literary Russian.

6 I translated into Russian Tsedryk's original Ukrainian example (8), which forces a clearer A-topical context.

7 In Russian, logophoric features (licensing 1st and 2nd person subjects) are easily recoverable from the situational context, which is characteristic of other partial NS-languages as well (Standard Finnish, Hebrew, Germanic; cf. Holmberg 2005; Vainikka & Levy 1999, among others). Many authors attribute this property to Sigurðsson's (2011) insight that every clause has available logophoric features in its C-domain, so that the speaker and addressee are always available as local antecedents (cf. Frascarelli 2007; Holmberg, Nayudu & Sheehan 2009). For the purposes of this chapter, I will lump together null G-Topics and null Topics related to logophoric features as representing those environments in which NSs are available in a "direct" way in Russian, unlike

speaker implies that both participants in the conversation know that the father is or has been sick. Thus, the answer contains given information and its subject can be dropped.

In (1c), however, the subject is interpreted as a sentence Topic, also known as Aboutness-Shift Topic (A-Topic), (see Frascarelli 2007, 2018, inspired by Reinhart 1981).[8] Here, even if the entity 'Ivan' is overtly present in the previous context, the second speaker turns to a new propositional content (he gives new information about Ivan): thus, the pronominal subject qualifies as an A-Topic, and must be now overtly realized (see Tsedryk, 2022).

There is a second context in which NSs in root contexts are available in Russian, which are not interpreted as G-Topics, and which do not necessarily involve logophoric features: non-emphatic subjects, coreferential with a nominative antecedent in the previous clause, i.e. successive instances of a subject already introduced in the discourse.

(2) *My*$_i$ *idëm*   *na ozero.*   ___$_i$ *nadeemsja tam*   *vstretit' Ivana*$_j$.
we.NOM   go.1PL   to   lake   (we) hope.1PL   there see.INF   Ivan.ACC
*(On$_j$)   *nam*   *obeščal*   *peredat' ključi.*
he.NOM   we.DAT   promised.M   pass.INF   keys.ACC
'We are going to the lake. We hope to see Ivan there. He has promised us to pass the keys.' (Tsedryk, this volume, (11))

In example (2), the second pronominal subject can be null, as it just follows up the information given in the first sentence (the reason why we are going to the lake). It is the second realization of a previously introduced identical subject (*my* 'we') that Frascarelli (2007) calls "continuing topic" in a "topic chain".[9] Successive realizations of a subject of this kind do not introduce new propositional content and thus cannot qualify as an A-Topic. The NS in (2) contrasts with the third pronominal subject in (2), *on* 'he', which qualifies as an A-Topic, i.e. it shifts the propositional content introducing new information about Ivan, which is

---

successive instances of A-topics, which need additional structure mediating between the antecedent and the NS. In the sections to follow, it will be shown that they need to form a "chain of topics" subject to certain grammatical requirements.

8 A(boutness-Shift)-Topics reintroduce a topic in the discourse or shift to a new propositional content without losing or replacing information from the previous context.

9 Tsedryk (2022) lumps together these two types of null Topics and identifies successive copies of A-Topics with G-Topics. In any event, root NS in Modern Russian must be characterized as a specific type of null Topic.

unrelated to the previous one. Thus, it must be overt (*on* 'he'), even if it refers to an element overtly present in the immediately previous sentence (*Ivan*).

As I noted in the initial remarks, I will follow Holmberg et al. (2009) in that T lacks an unvalued D-feature in partial NS languages, so T cannot "mediate" in the transmission of a referential index to a minimal φP. However, referential indexes can be transmitted in a topic chain, just by index identification between the two links of the chain (Holmberg et al. 2009). According to Frascarelli (2017), in a topic chain the φP copies the referential index of the D-valued antecedent through a null operator in CP, probably through a dedicated ShiftP, obtaining a referential interpretation. If this is the case, the φP does not need to be overtly realized (i.e. it can be realized as NS), because the topic chain can undergo chain reduction (Bobaljik 2002), and only the highest link of the chain needs to be pronounced.

(3) $[_{DP_i} \ldots [_{CP} [_{ShiftP} Op_{[+aboutness]} [_{TP} \varphi P_i\ T \ldots]]]]$

It is worth noting that Russian allows null coreferential pronominal subjects as successive realizations of an A-Topic only if there is no potential intervener blocking the relation between the NS and its antecedent in the topic chain (cf. Tsedryk 2012, 2015; Livitz 2014). The presence of an overt intervener between the C-layer and a subject forces the overt realization of the subject (4b), which can otherwise be null, as in (4a):

(4)   a.   *Ja$_i$ tol'ko čto   videl   Svetu$_j$.*   (*Ona$_j$*)   *skazala   mne$_i$,   čto   naš*
           I   just   saw   Sveta.ACC   she.NOM   said.F   I.DAT   that   our
           dom   uže   prodan.
           house   already   sold
           'I've just seen Sveta. (She) told me that our house had already been sold.'

      b.   *Ja$_i$ tol'ko čto   videl   Svetu$_j$.*   **Mne$_i$**   *(ona$_j$)   skazala,   čto   naš*
           I   just   saw   Sveta.ACC   I.DAT   she.NOM   said.F   that   our
           dom   uže   prodan.
           house   already   sold
           'I've just seen Sveta. She told me that our house had already been sold.'
           (Tsedyk 2015: 349)

As for silent G-Topics / Logophoric Topics, I will adopt Sigurðsson's (2011) account of Germanic null Topics. He argues that every referential definite argument, whether silent or not, matches an "edge" feature available in C. The matching is performed in a direct way, i.e. with no operator or extra-structure mediating the

relation (cf. Tsedryk 2015 specifically for Russian).[10] This is an alternative way for NSs in Russian to receive a referential index and, thus, be interpreted as referential, albeit null, at PF.

(5) $[_{CP} \text{Op}_{[+edge]} [_{TP} \varphi P_i \text{ T}. . .]]]$

The matching illustrated in (5) is possible only if the NS has direct access to the empty [Spec,C], i.e. when there is no other overt D-element intervening between C and the φP. In (6), the presence of a left-dislocated direct object (*Svetu*) does not prevent the NS from checking its φ-features in a regular way (with T), as reflected in the verbal form (here, plural). However, the left-dislocated object acts as an intervener between the logophoric edge-feature contained in C and the NS. As a consequence, the NS in (6) gets an arbitrary reading and cannot be interpreted as referential:[11]

(6)   *Svetu* ___ *nedavno videli na rynke.*
       Sveta.ACC   recently saw.PL at market
       'Sveta was seen recently at the market.' / '*We saw Sveta recently at the market.' (Tsedryk 2015: 349)

2.2 Embedded contexts and "control"

In the previous section I presented two types of null Topics available in Modern Russian, which can be realized as NSs (Logophoric Topics / G-Topics and successive instances of A-Topics in topic chains). Now I will turn to a third type of NSs in Modern Russian available in embedded contexts, i.e. NSs c-commanded by a close antecedent.

In Modern Russian, a NS with individual reference can be inserted in a subordinate clause without being committed to a topic interpretation, as in (7) below. This property is shared by Russian and other partial NS languages, e.g. Finnish,

---

10 Edge features are related to various interpretations: (i) a logophoric feature anchors the interpretation of the NS in the speaker-hearer domain, or (ii) a "topic" feature, in Sigurðsson's (2011) sense, is taken here to be a situational or contextual deictic feature, resulting in a NS interpreted as a G-Topic.

11 In this respect, the crucial difference between Russian and Germanic is that Modern Russian is not a V2 language and does not have V-to-T movement (Bailyn 2012, Gribanova 2013), so V remains low, and the number of potential interveners is more reduced in Russian than in Germanic. For example, pre-verbal adverbials and PPs do not behave as interveners (Tsedryk 2015), as opposed to what happens in Icelandic (Sigurðsson 2011).

Brazilian Portuguese, Hebrew, and Marathi (cf. Landau 2004; Holmberg et al. 2009; Modesto 2011).

(7) *Vrač*ᵢ *skazal, čto* ___ᵢ *primet* *bol'nyx*ⱼ.
doctor said that (he) will receive.3SG sick people.ACC
'The doctor said that he will see patients.'

In general terms, embedded NSs in Modern Russian are partially conditioned by pragmatic considerations: an embedded subject can be either overt or null if it implies G-topicality (8b), whereas it is preferably null otherwise (8a). Let us consider the following example provided by Pavel Graschenkov (personal communication).

   In (8), the participants are planning a party. If there is no previous expectation (positive or negative) about Petya's coming to the party and someone unexpectedly utters (8a), the presence of a coreferential NS in embedded position is highly preferable, at least in colloquial language. However, someone can introduce Petya in the discourse by asking a question like 'What about Petya?' and, thus, enforcing some previous expectation about Petya's coming. In this case, the embedded clause can function as given, and its subject can be null or overt (a G-Topic), as in (8b).

(8) a. *Petja*ᵢ *skazal,* [*čto* ?*on*ᵢ / ___ᵢ *pridët*].   (unexpected: no
      Petja said that he.NOM will come.3SG topicalization)
      'Petja said that he would come.'
   b. *Petja*ᵢ *skazal,* [*čto* *on*ᵢ / ___ᵢ *pridët*].   (embedded G-Topic)
      Petja said that he.NOM will_come.3SG
      'Petja said that he would come.'

The pragmatic situation of the latter context (8b) is equal to that of G-Topics in root clauses (e.g. in example (1b) above): the subject will be preferably null if the participants in the conversation expect that Petya is coming, whereas it will be preferably overt if the participants do not expect him to be coming (Egor Tsedryk, personal communication). In all likelihood, embedded G-Topics have access to the same edge features as root G-Topics, and the mechanism of licensing null G-Topics must be the same in both contexts.

   This is consistent with the idea, put forward in section 2.1, that null G-Topics are quite freely licensed in Modern Russian, both in root and embedded contexts, given the right pragmatic conditions. However, there is no commitment to a G-Topic interpretation of the embedded clauses in (7) and (8a), which requires a different analysis of the corresponding NSs.

Previous work on this kind of NSs has established the existence of common properties in several partial NS languages. More specifically, as Landau (2004) shows for Hebrew, Rodrigues (2004) and Boeckx, Hornstein & Nunes (2010) for Brazilian Portuguese, and Tsedryk (2012); Livitz (2014: 71ff, 110ff) for Russian, embedded NSs in finite clauses exhibit the following properties: (i) need for a local c-commanding antecedent; (ii) ban on split antecedents; (iii) sloppy interpretation under ellipsis; and (iv) *de se* reading.[12]

These are some of the typical characteristics of what is called "obligatory control" in the generative tradition, originally proposed for PRO in non-finite clauses (see Chomsky 1981; and Hornstein 1999). The similarities between the classic PRO and NSs in finite embedded clauses in partial NS languages led some authors to use the term "finite control" for this phenomenon (Boeckx, Hornstein, & Nunes 2010; Ferreira 2009; Modesto 2011; Landau 2006, among others). I will avoid this term and the related discussion as it is irrelevant for the purposes of this chapter, and stick to the main idea that embedded NSs are necessarily c-commanded by an antecedent in Modern Russian, as shown in (9).[13]

(9) a. *Petja*$_i$ *ne znaet, kak on*$_{i/j}$ / ___$_{i /*j}$ *sjuda popal.*
    Petya not knows how he.NOM   here ended_up.M
    'Petya doesn't know how he ended up here.'
    (Livitz 2014: 84)

  b. [*Doč'   prezidenta*$_j$]$_i$ *ob'javila, čto ona*$_i$ / ___$_{i /*j}$
    daughter president.GEN announced that she.NOM / (she.NOM)
    vystupit   s   dokladom.
    will_perform.3SG with speech
    'The president's daughter announced that she will give a speech.'
    (Livitz 2014: 72)

The example in (9a) shows that an embedded NS must have a coreferent antecedent in order to be licensed; otherwise, it must be realized as an overt pronoun. Example (9b) illustrates the fact that the antecedent of an embedded NS must c-command the NS; if not, the embedded pronoun must be overt.

---

12 These properties are not all met in other partial NS languages, namely in Finnish and Marathi (Holmberg et al. 2009), but they are fulfilled in the case of Russian. The linguistic productions in Russian that could be taken as potential exceptions are convincingly discarded by Livitz (2014), who ascribes them to independent processes.

13 For the other properties of "finite control" in Russian, as well as their parallel in non-finite clauses, see a detailed account and abundant examples in Livitz (2014).

In Modern Russian, the interpretation of the reference of an embedded NS is carried out under strict locality (Tsedryk 2012; Livitz 2014; Shushurin 2017; Madariaga 2018), at least whenever they are not interpreted as G-Topics. This is shown in (10a), in which the only possible antecedent for the NS is the closest c-commanding antecedent. Strict locality is also the reason why a plural intervener renders ungrammatical the sentence in (10b), in which the embedded verb is singular. [14]

(10) a. *Maša*$_i$ *boitsja,* *čto* *Anja*$_j$ *dumaet,* *čto* ___$_{j/*i/*k}$ *ne*
Masha fears that Anya thinks that not
pridët *vovremja.*
will come.3SG on time
'Masha is afraid that Anya thinks that (Anya/*Masha) won't arrive on time'

b. **Majja*$_i$ *boitsja,* *čto* *roditeli*$_j$ *dumajut,* *čto* ___$_i$ *ne*
Maia fears that parents.PL think that not
pridët *vovremja.*
will come.3SG on time
Intended: 'Maia is afraid that her parents think that (Maia) won't arrive on time'

As for the mechanism underlying licensing of this kind of embedded NSs, previous accounts in the literature have adopted different views. Tsedryk (2012) and Madariaga (2018) adopt the movement theory of control proposed by Boeckx, Hornstein & Nunes (2010) for Brazilian Portuguese, identifying almost completely finite and non-finite control. Other works, such as Landau (2004); Livitz (2014); and Tsedryk (2015), assume an Agree-based analysis of control.

For the purposes of this chapter, I will follow Landau's (2015) account, adapting it in the case of finite control. Technical details aside, Landau (2015) describes the mechanism of "obligatory control" in complement infinitive clauses in the following way: the matrix controller binds a variable at CP, which, in turn, binds a minimal pronoun with unvalued φ-features in the complement clause (PRO). A pronoun which has valued its φ-features and is bound by a matrix antecedent can be realized as null (PRO) at PF, as far as it shares its reference with that antecedent. Interestingly for us, according to Landau (2015), Agree as such

---

14 Incidentally, NSs in embedded finite contexts in partial NS languages differ from their equivalents in consistent NS languages in that the latter do not undergo locality restrictions of this kind (cf. Alonso-Ovalle et al. 2002 for Spanish):

(i) *Pedro*$_i$ *le* *ha* *dicho* *a* *Juan*$_j$ *que* ___$_{i/j/k}$ *no* *llegará* *a* *tiempo.*
Pedro CL has said to Juan that not will arrive on time
'Pedro told Juan that he (=Pedro, Juan, or someone else) will not arrive on time'.

plays no role in establishing the semantic antecedence relation; it just matches the φ-values of PRO. The reference of PRO is established in the relation with the binding antecedent through an intermediate variable in CP.[15]

On the other hand, Landau (2015) defines as Non-Control all those instances in which T is both tensed and inflected for φ-features. Under this view, Russian embedded finite clauses with NSs would not be instances of control. However, given the common properties of finite embedded NSs and infinitive PRO in Modern Russian, I will assume, as in the case of root NSs, that the NS is a minimal φP which values its φ-features with T in a canonical way, thereby resulting in verbal agreement. Yet, as in root finite clauses, T lacks an unvalued D-feature, so the φP needs to get its reference in an independent way, i.e. from the closest c-commanding (compatible) antecedent. If this is the case, then recoverability of the reference of the embedded pronoun is again the reason why it can be realized as null at PF. This mechanism is a simpler instance of variable binding than control in non-finite clauses in that φ-features of the minimal φP are valued straightforwardly within the embedded clause, but we still need to posit a bound variable "mediating" between the controller and the embedded φP, as Landau (2015) does in the case of infinitive control.

(11) $[DP_i \ldots [_{CP} Op_i C [_{TP} \varphi P_i T \ldots]]]$

Notice that the mechanism of licensing embedding NSs of this kind is different from topic chains (cf. Section 2.1) because there is no pragmatic or informational commitment to A-topicalization here, and a Shift projection does not make any sense in this case. The different nature of topic chains and embedded finite NSs is evidenced by several properties, such as the ability of antecedents in topic chains to bind a wider range of subjects (e.g. dative subjects) than in embedded finite NS structures (Egor Tsedryk, personal communication).

Another remarkable difference between topic chains and embedded NSs is the fact that the interveners blocking NSs in topic chains are different from those blocking NSs in finite embedded contexts. In Section 2.1 I showed that any

---

15 As for φ-feature valuation in non-finite control, Livitz (2014) specifies that the embedded T, which is the first potential Probe for PRO, has unvalued φ-features itself and cannot match PRO's features. On the other hand, the inflected matrix T matches its unvalued φ-features with the matrix subject in a regular way, obtaining the corresponding φ-feature values. Then, this matrix T is able to probe further into the embedded clause, matching the features of PRO as well, thus creating a sort of Agree chain. In order to avoid a violation of the Phase Impenetrability Condition, Livitz (2014: 51ff) argues further that, in fact, the Agree relation between PRO and its antecedent is mediated by the matrix little v, endowed with unvalued φ-features. More specifically, when little v cannot value its features within its c-command domain (the complement CP), it probes upward and ends up with the features of the matrix subject.

referential overt entity (arguments, obliques, etc.) can act as an intervener in a topic chain, preventing φP from copying the referential index of its antecedent through the null operator in CP. In example (4b) above, for instance, a left-dislocated dative goal (*mne* 'me') blocks Topic drop.

In contrast, a left-dislocated dative goal does not inhibit subject drop in finite embedded clauses (12a). Only subject-like elements located in the EPP position (e.g., a dative subject) can act as interveners in this case (12b), whereas subject drop can be performed if the EPP position is available (12c).

(12) a. *Maša*$_i$ *dumaet,* *čto* **mne**   ___$_i$ *skazala* *nepravdu.*
          Maša   thinks   that me.DAT   said   non-truth
          'Masha thinks that she told me a lie.'

      b. ?*Petja$_i$ *skazal,* *čto* **Maše**$_j$ ___$_i$ *nravitsja.*
          Petja.NOM   said   that Maša.DAT   pleases

      c. *Petja*$_i$ *skazal,* *čto* ___$_i$ *nravitsja* *Maše*$_j$.
          Petja.NOM   said   that   pleases Maša.DAT
          'Petja said that Masha likes him.'
          (Livitz 2014: 111–112)

Even though word order in Russian is quite free, there is a simple test to check the position of the dative element in these sentences. Following Bailyn (2004, 2012), I assume that word order rearrangements in Russian are due to one of the two following basic mechanisms at the clausal level, i.e. Generalized Inversion or Dislocation. Generalized Inversion is an EPP-driven A-movement, whereby any XP can be raised to [Spec,TP] position and satisfy the EPP.[16] In case a non-nominative XP raises to [Spec,TP], the verb raises to T over the subject (unlike in regular SVO sentences in which V goes only as far as v). As a result, this mechanism typically renders OVS/XVS word orders. Dislocation, on the contrary, is a discourse-driven A'-movement, by virtue of which any XP can be dislocated to the left periphery, to check some informational feature in the CP domain.

In our examples, (12a) is not an instance of Generalized Inversion. It does not display the canonical OVS/XVS order (e.g. . . ., *čto nepravdu skazala mne ona* / . . ., *čto nepravdu mne skazala ona*) and, as expected, it becomes ungrammatical once the pronominal subject is silenced (. . ., **čto nepravdu skazala mne* ___$_i$ / . . .; *čto

---

16 I assume that V-to-T movement satisfies the EPP in consistent Null Subject Languages (Alexiadou & Anagnostopoulou 1998, among others), whereas in partial Null Subject Languages, the EPP is canonically satisfied by some XP located in [Spec,TP] or by a referential / definite pronoun, which can be eventually realized as null (Holmberg, Nayudu & Sheehan 2009, Roberts 2009, among others).

nepravdu mne skazala ___ᵢ). So (12a) can be either (i) an instance of left-disloca-
tion (when *mne* is interpreted/pronounced as a focus or a contrastive Topic, 'it is
me that she told the truth'), or (ii) VP/vP-level Inversion, typical of goals (Bailyn
2012: 311–312), when it is interpreted and pronounced as pragmatically neutral.
But in neither case does *mne* intervene between the embedded φP and its control-
ler, because it does not compete for the Topic position of the intermediate bound
variable or the embedded EPP position in (11).

In (12b), however, the embedded clause is a typical instance of XVS (General-
ized Inversion), in which the dative experiencer A-moves to [Spec,TP], as argued
in detail by Bailyn (2012: 161, ff.). From that position, the dative experiencer acts
as an intervener, blocking the drop of the pronominal embedded subject, while
it does not do so if it stays low in the structure (12c). The reason follows straight-
forwardly from the configuration in (11), in which the [Spec,TP] subject position
must be occupied by the minimal φP itself, in order to be high enough to enter
into a reference transmission relation with the matrix subject controller.

3 Old Russian NSs with individual reference (Grammar 1)

In this section, I show that NSs in Old Russian formed a homogeneous group,
as opposed to NSs in Modern Russian (see Section 2). Old Russian NSs were reg-
ularly licensed in the domain of a T whose unvalued D-feature was valued by a
null Topic at [Spec,CP], while focused or contrastive subjects had to be realized as
overt. In fact, if the subject was a pronoun, it *had* to be realized as null (meaning
that a Topic or a topic feature was always available at CP in Old Russian), unless
there was an additional requirement that forced its overt realization (namely a
[+focus] or [+contrastive] feature). This implies that A-Topics were also realized
as null in Old Russian, as we will see in what follows.

As far as we can judge from the texts, Old Russian was a "well-behaved"
consistent NS language (Meyer 2011; Eckhoff & Meyer 2011; Madariaga 2015,
2018; Jung 2018). In general terms, until the 15th century, the baseline realization
of pronominal subjects was null in informationally neutral (i.e. 'non-stressed')
positions in terms of Borkovskij (1949, 1978), meaning non-emphatic, non-
contrastive, non-focal positions (Meyer 2011).[17] Let us consider an example of

---

17 Overt pronominal subjects in Old Russian were restricted to 'stressed' contexts (information-
ally "emphatic," contrastive, focal), as in (i).

a series of 2nd person root NSs, similar to Modern Russian Logophoric Topics / G-Topics in (1a-b):

(13) *Počto* ____i *ideši opjatъ,* ____i *poimalъ esi vsju danь.*
     what for   go.2.SG again   took.M   AUX.2.SG whole   tax
     'Why did you come back? You collected the whole tribute already.'
     [*Laur. Chr.*, 14v]¹⁸

Despite potential similarities with Modern Russian in this specific context, NSs in Old Russian were not restricted to this configuration. They were licensed in a uniform way, regardless of the syntactic context (root or embedded), their topical status (G-Topic or A-Topic), and the specific nature of their antecedent, as a far as it was itself some kind of Topic.

As a matter of fact, pronominal subjects in Old Russian were dropped whenever they were informationally "neutral" and their reference recoverable from the context, regardless of other "aboutness shifts" in the discourse. That is, even the subjects that are nowadays interpreted as A-Topics, always overt in Modern Russian, were null in Old Russian (Eckhoff & Meyer 2011). Let us see a pair of examples extracted from the part of the *Laurentian Chronicle* that narrates Prince Oleg's life:

(14) a. *I viděvše že Grěcě*_i *ubojašasja* (. . .). *I ustavi*
        and seeing part Greeks feared   and stopped.3SG
        *Olegъ*_j *boi*_k. *I* ____i *vynesoša emu*_j *brašna i*
        Oleg.NOM soldiers and   brought.3PL him food   and
        *vino*_m *i* ____j *ne prija ego*_m ____m *bě bo ustroeno*
        wine and   not took.3SG it   was part prepared
        s otravoju.
        with poison
        'When they saw that, the Greeks became afraid (and asked Oleg to stop the war). Oleg stopped his soldiers. And [the Greeks] brought him food and wine, and [Oleg] did not drink it, because [the wine] was poisoned.' [*Laur. Chr.*, 15R]

(i) *Ne **jazъ** bo počalъ bratъju biti no **onъ**.*
    not I.NOM PART started.M brothers hit but that.NOM
    'It was not me, but him, who started to attack our brothers.' [*Laur. Chr.*, 24]

18 Old Russian examples have been extracted from conversational-like and narrative passages of the 14th-century Laurentian Chronicle (Karskij 2001 [1926–28]). The examples extracted from the Radziwill codex (15th century) are marked with an "R".

b. *I*   *(Olegъ$_i$) s'sěde*   *c*   *konja* (. . .) *i*   ___$_i$ *vъstupi*
and (Oleg) got_off.3sg from horse   and   stepped.3sg
nogoju   *na lobъ. I*   *vynuknuvši zmia$_j$*   *zo*   *lba*
with_foot on skull and getting.out snake.nom from skull
i$_i$ *ukljunu v*   *nogu i*   *s*   *tog(o)* ___$_i$ *razbolěs(ja)*
him.acc bit.3sg on foot and from this   got sick.3sg
i   *umre.*
and died.3sg
'And [Oleg] got off his horse (. . .) and stepped on the skull. And a snake came out of the skull and bit his foot, and [Oleg] got sick from this bite and died.' [*Laur. Chr.*, 19R–19Rv]

Here is the context for example (14a): Oleg makes constant incursions into Byzantium and the Greeks pretend to agree to pay Oleg a tax if he stops attacking them. They set up a welcome reception for him, but their real intention is to poison him. In example (14a), we observe the first topic shift from Oleg (and his soldiers) into a NS referring to the Greeks, who bring Prince Oleg food and poisoned wine. Then, there is the second shift back into a new NS retaking a reference to Oleg. In between, there is a further shift into a NS referring to the wine, which had appeared before. All the three NSs, clearly qualifying as A-Topics, would have to be overtly realized in Modern Russian.

As for the example in (14b), the magic men prophesize that Oleg will die because of his favorite horse. As a result, Oleg decides not to ride it any more. The horse finally dies. Oleg comes to see the bones of the dead horse, and is happy that the prophecy has not been fulfilled. But he steps on the horse's skull and gets bitten by a snake that comes out from the skull, and the bite leads to his death. In example (14b), we observe first an A-Topic shift from Oleg into the snake, but this intermediate A-Topical subject (the snake) does not prevent the author from referring to Oleg again later with the help of a NS.

This is a recurrent pattern in the old chronicles, whereas such root NSs would be ungrammatical in Modern Russian. Dropping coreferential subjects, regardless of their topical nature, implies that the specific type of topicalization, as well as the degree of embedding, is not relevant in licensing NSs in Old Russian.[19]

---

19 I do not overlook the fact that overt φ-marking in morphology is crucial for the right interpretation of NSs in consistent NS languages and sometimes, though not always (cf. fn. 14), enforces overt realization of the pronoun in case of ambiguity. In (14a), the presence of number morphology on the verbs unequivocally points to the right interpretation of the reference of the NSs, while this is not the case in (14b). On the other hand, overt φ-marking, thought impoverished with respect to Old Russian, is present in all verbal forms in Modern Russian (gender and

Old Russian is not the only consistent NS language functioning in this way. Jiménez-Fernández (2016) shows that Spanish is very permissive with regard to the referential identification of NSs. For instance, successive Topics in a topic chain are licensed by any topical antecedent, whatever its nature (A-Topics, G-Topics, or even contrastive Topics), as far as they convey the [+given] feature, meaning that their reference is mentioned in (or inferred from) the previous context or shared knowledge of the situation. The fact that a pronominal subject has to be realized as null, in absence of a [+focus] or [+contrastive] feature, implies that a topic feature (no matter which one) is always available at CP in Spanish. A similar picture emerges in the case of Old Russian.

From this point of view, we do not need to "dissect" the left-periphery in specific heads in order to account for NSs in Old Russian, while, paradoxically, we must consider the possibility of distinguishing at least some dedicated position for A-Topics in Modern Russian (in contrast to G-Topics) as relevant for the distinction between null vs. overt pronominal subjects (see Section 2.1).[20]

As I explained in Section 1, for consistent NS languages, I follow Frascarelli (2007) and Holmberg & al. (2009) in stating that T is endowed with an unvalued D-feature, which has to be valued. The reference of a NS is specified in an indirect way: the referential index of a null Topic in [Spec,CP] is copied by the unvalued D-feature of T, and, finally, through Agree, by the φP (=NS), which at the same time matches its φ-features, resulting in rich verbal agreement.[21]

(15) $[_{CP} \text{Topic}_i \text{ C } [_{TP} \text{φP}_i \text{ T}_{D:\_} [_{VP} \text{V} . . .]]]$

---

number for past forms, person and number for non-past forms), but there is still a clear contrast in the licensing of NSs between the two varieties.

20 Incidentally, a short survey that I conducted among the speakers of standard peninsular Spanish reveals that pronominal topical subjects are null, regardless of A-Topic shifts. For space reasons, I just give a brief example in (i), in which the A-Topic 'Pedro' shifts back at a certain point to the initial subject 'Juan' and, still, we do not need (and the speakers surveyed do not accept) an overt pronoun to recover Juan's reference.

(i)   *Juan$_i$ se había tomado un café y ___$_i$ se iba a tomar otro, pero entonces vino Pedro$_j$ y le$_i$ dijo que olía muy bien a café, que a ver si ___$_j$ podía tomarse uno. Y como no quedaba más, ___$_i$ / *$^{??}$él$_i$ decidió darle$_j$ el suyo.*
      'Juan$_i$ had some coffee and ___$_i$ was going to have some more, but Pedro$_j$ came in and told him$_i$ that there was a nice smell of coffee, that ___$_j$ would like to have some. And, as there was no more left, ___ / *$^{??}$he$_i$ (=Juan) decided to give him$_j$ his own coffee.'

21 Holmberg et al. (2009: 75ff) argue that the D-feature of T and φ-feature matching on the NS are intimately related for the following reason: as long as the D-feature of T has no morphological expression of its own, the only way it can be spelled out is by spelling out the person and number features, which is typically found in consistent NS languages with rich verbal morphology.

I assume that this is the basic mechanism of licensing NSs in Old Russian, too. As a complementary condition, Old Russian displayed V-to-T movement, evidenced by the raising of overt verbal auxiliaries and the position of clitics (Jung 2018) – see examples (13) and (16a) below. V-to-T movement was crucial for licensing NSs in Old Russian, as it ensured that T checked the EPP and, after valuing its D-feature, was able to transmit the corresponding reference to the NS together with Agree. (Recall that, in Modern Russian, T lacks a D-feature, V does not raise, and T does not take part in the referential index transmission to the NS).

As for embedded NSs in Old Russian, as compared to Modern Russian, they did not undergo control requirements in most contexts in either finite (16a) or infinitive clauses (16b).

(16) a. *Kde   es(tъ) konь*$_i$ *mъi*$_j$ *jegože*$_i$ ___$_j$ *bě*
          where   is   horse.NOM my.NOM which.ACC   AUX.1.SG
          postavilъ kormiti i   bljusti   jego$_i$?
          put.M.SG   feed   and   take care him.ACC
          '(Oleg said) Where is my horse, whom I ordered to feed and take care of?'
          [*Laur. Chr.*, 19R]

      b. *Molisja* [*za mja*$_i$]   *otče   čestnyi* [___$_i$ *izbavlenu*$_i$ *byti*
          pray   for me.ACC   father dear   released.DAT   be.INF
          ot   seti   neprijazniny].
          from   this   devilment
          'Honorable Father, pray for me to be saved (me) from devilment.'
          [*Laur. Chr.*, 71v]

In (16), I illustrate two instances of embedded NSs which lack a local c-commanding antecedent and would require an overt pronoun in Modern Russian. Example (16a) includes a finite relative embedded clause. Here, the embedded subject refers to the first person and is realized as null, even if the only potential antecedent in the matrix clause is the possessive *moi* 'mine', which is not in a c-commanding position (cf. (9) and (10) above in Modern Russian, in which the NS must be bound by the matrix subject). Example (16b) illustrates a non-finite embedded clause and, again, the semantic antecedent of the NS is not a proper controller because it is located within a PP (*za mja* 'for me').

The parallelism between finite and non-finite embedded clauses in Old Russian was not restricted to their common lack of a local antecedent to license embedded NSs (see Madariaga 2015, 2018 for a detailed account). In Old Russian, overt dative subjects in embedded infinitive clauses depending on declarative and desiderative verbs (17) were almost as common as nominative subjects in finite clauses.

(17)   *Ty*ᵢ   [*so*   *mnoju*ⱼ]   *cělovalъ*   *kr(e)stъ*   [*xoditi* **nama**ᵢ₊ⱼ *po*   *odinoj*
      you   with   me   kissed.M.SG   cross   go.INF   we.DAT   by   one
      dumě   *oběma*ᵢ₊ⱼ].
      decision   both.DAT
      'You and me swore (lit. kissed the cross) that we both will do it the same
      way.' [*Laur. Chr.*, 170v]

Example (17) illustrates an embedded infinitive clause, the complement of a declarative verb *cělovatь krestъ* 'to swear', containing an overt dative pronominal (emphatic) subject *nam* 'us', coreferential with two split antecedents.

     There are various differences between embedded infinitive clauses in Old Russian and Modern Russian; the most relevant one for the purposes of this chapter is that an embedded clause, the complement to a volitive matrix verb with non-coreferential subjects, as in (16b), cannot be realized as an infinitive clause nowadays; an obviation finite structure (with *čtoby*) is used instead (18).[22]

(18)   *Molis',*   *čtoby*   *ja*   *byl*   *izbavlen*   *ot*   *neprijatnostej.*
      pray   so that   I.NOM   be.M.SG   released.M.SG   from   problems
      'Pray, so that I am saved from the problems.'

Regardless of the widespread lack of control in finite and non-finite clauses in Old Russian, there are certain configurations which suggest that obligatory control could exist in Old Russian. Namely, some instances of embedded infinitive clauses that performed as complements of volitive verbs with coreferential subjects, such as (19), were indistinguishable from control structures of the Modern Russian type (Madariaga 2011).[23]

---

22 Infinitival clausal complements of declarative verbs, like the one in (17), were lost in favor of finite CPs introduced by the complementizer *čto*.

23 Infinitive clauses headed by a non-coreferential overt dative subject, complements to volitive verbs, albeit rare, are found in older Russian texts. In even older stages, i.e. Old Church Slavonic, a volitive verb could sometimes take an infinitive clause headed by an overt coreferential subject (i):

(i)   ___ᵢ   *moljaaxǫ*   *i*ⱼ   [*prijti*   **jemu**ⱼ].
      beg.3PL   him.ACC   come.INF   he.DAT
      'They begged him to come.' (Old Church Slavonic: *Codex Suprasliensis* 16, 103v)

(19) *Egdaž(e) trebuetъ na voinu iti, sii*ᵢ *xotjat(ъ)* [___ᵢ
 if is needed to war go these.NOM want.3PL
 počti c(ěza)rja vaše(go)] (. . .) *da budutъ.*
 honour.INF tsar.ACC yours.ACC let be.3PL
 'If you need to gather an army (lit. to go to war), and these (=the Russians)
 want to join your king (. . .), so be it.'
 [*Laur. Chr.*, 18R]

The similarity between finite and non-finite embedded clauses, as well as the existence of at least some structures that could be interpreted as "controlled" in Old Russian, will be significant for the diachronic development of embedded NSs in the new system, to be explained in section 4.

4 Shifting from Grammar 1 into Grammar 2

Following the mainstream in formal accounts of diachronic syntax, I follow Lightfoot (1979, 1999) in stating that language change is driven by the abductive reanalysis of the input the learner receives. The input can become slightly modified because of some previous related or unrelated change, and this gives the possibility for some learners to analyze the data in a new way, giving rise to a structure that is different from that which previous generations had acquired.[24] In this section, I explain the diachronic development from the Old Russian system of NSs (Grammar 1) into the new system found from Late Middle-Modern Russian (Grammar 2) on.

This could not be a one-to-one change, as we go from a homogeneous mechanism of licensing NSs in Old Russian into a heterogeneous system of different mechanisms, depending on the specific configuration of the NS in later Russian. Before elaborating on the details of the change, I schematize the relevant structures in (20):

---

24 I will assume a cue-based approach in a broad sense: learners detect in the input the relevant portion of structure to set a parameter, or to posit a structure or a feature (Lightfoot 1999).

(20) Shift from Grammar 1 to Grammar 2

Grammar 1		Grammar 2
$[_{CP}$ Topic$_i$ C $[_{TP}$ φP$_i$ T$_{[D:\_]}$ $[_{VP}$ V. . .]]]	>	$[_{CP}$ Op$_{[+edge]}$ C $[_{TP}$ φP$_i$ T. . .]]] (G-Topics)
		$[$DP$_i$ . . .$[_{ShiftP}$ Op$_{[+aboutness]}$ $[_{TP}$ φP$_i$ T. . .]]] (topic chains)
		$[$DP$_i$ . . . $[_{CP}$ Op$_i$ C $[_{TP}$ φP$_i$ T. . .]]] (embedded NSs)

The historical development we observe in the texts evidences a series of changes that ended up in the loss of the consistent NS status of Russian (Borkovskij 1949, 1978; Ivanov 1990; Jung 2018; Kibrik 2013; Meyer 2011; Migdalski 2013; Zaliznjak 2004, 2008, among others). The process was complex and took at least five centuries to complete.

Meyer (2011) observes that non-emphatic subjects were null until the 15th century, but then, between the 16th and 17th century, 1st and 2nd person weak pronouns experienced a notable increase. According to Zaliznjak (2004), by the 15th century, verbal auxiliaries started to be generated in a lower position, and verbs remained lower than raised pronominal elements (clitics). In other words, V-to-T movement was lost at that time. In formal terms, this change sparked a whole range of consequences. For instance, T could no longer check the EPP by itself, and pronominal clitics and verbal clitic auxiliaries began to disappear. Another important consequence was that the input that learners received became slightly modified and gave rise to a change in the *pro*-drop character of the language too.

Once V remained in vP, learners did not receive the relevant cue to posit a D-feature in T any more, because V's low position preempted the establishment of the direct syntactic relation between T and C the way it had proceeded before. The loss of the D-feature on T was underpinned by the loss of (fully inflected) verbal clitic auxiliaries in the language (analyzed in detail in Zaliznjak 2008), which had formerly strengthened the acquisition of T as a head clearly hosting φ-features and playing a role in D-feature transmission together with Agree.

As a consequence, pronominal subjects were not able to receive a referential index together with D-feature valuation of T any more and started to be overtly realized. Otherwise, they were interpreted as non-referential (generic, arbitrary. . .). This change is pinpointed in the texts by the sudden rise of overt pronouns of 1st and 2nd person, which had formerly functioned only as emphatic pronouns (focused or contrastive), and now spread as overt pronouns in non-emphatic positions, as well (Borkovskij 1978; Ivanov 1990; Meyer 2011). Dedicated 3rd

person pronouns did not exist and other pronouns (mainly demonstratives) were used instead for emphatic purposes.[25]

Looking at the contrast between both stages of Russian, we need to explain why it is possible that a homogenous mechanism of licensing NSs in Grammar 1 yielded at least three different ways to license them in Grammar 2, as represented in (20). I will argue that Russian became a "non-*pro*-drop language", but rapidly (or maybe in parallel) evolved into a "partial NS language" by reanalyzing residual subject gaps that were persistently present in the learners' input. In the absence of a unique way of licensing NSs (due to the previous changes in T and NSs), learners reanalyzed these residual gaps in the best way they could, according to the conditions available in each specific context.

In Grammar 2, when V-to-T was lost, together with T's D-linking ability, pronominal subjects had to be overtly realized in order to be interpreted as referential. In other words, Russian became factually a non-*pro*-drop language. However, learners of Grammar 2 would still receive referential subject gaps in their inputs, which corresponded to (i) NSs generated by speakers of Grammar 1, i.e. older generations of speakers; and (ii) some instances of PRO, i.e. subject gaps in control infinitive clauses complement to volitive verbs (examples like 19 above).

As is well known in historical linguistics, residual structures and elements after diachronic change can experience different outcomes. Specifically, (i) they can just die out as time goes by, or (ii) be "recycled" or reused with a new value, i.e. further reanalyzed to fulfill a new function. I will argue that residual referential NSs from Old Russian followed this second path.

Learners of Grammar 2 could not interpret subject gaps in the old fashion, that is as referential, in absence of a D-linking T in the language. They regularly interpreted them as generic, arbitrary, and so on. Nonetheless, learners found alternative ways of acquiring at least some of those referential subject gaps they received, namely by "delving" further in the structure for a proper mechanism of index transmission. Sometimes, they found an edge/topic feature at CP that

---

25 Zaliznjak (2008) and Meyer (2011) show that 3rd person auxiliaries were lost very early (maybe prehistorically), while 3rd person non-emphatic overt pronouns spread much later, by the 16th –17th centuries. Meyer (2011) argues that the lack of agreement itself marked 3rd person in Old Russian. The further extension of 1st/2nd person pronouns from the 15th century on probably helped reanalyze the former demonstrative pronoun *onъ* 'that' as a 3rd person pronoun, which spread automatically as a non-emphatic pronoun, together with the other persons. This reanalysis can be naturally explained by Input Generalization of the person feature, i.e. the extension of the syntactic behavior affecting 1st and 2nd person also to 3rd person. Input Generalization is a general optimization principle operating in acquisition / diachrony, defined by Biberauer & Roberts (2017: 147ff) in the following way: "[i]f a functional head sets parameter p_j to value v_i, then there is a preference for similar functional heads to set p_j to value v_i."

could transmit a referential index to the NS from a given Topic, which was easily recoverable on the basis of contextual shared information, giving rise to null G-Topics and null Logophoric Topics. On other occasions, learners could recover the reference of a NS by relating it to a higher referential antecedent through CP, whether a null A-Topic in a topic chain (in root clauses) or an embedded subject bound by an antecedent in the matrix clause ("control").

In the last two cases, the immediate consequence was the emergence of previously nonexistent locality restrictions. In topic chains, the lower null Topic must be identical to the other copies in the chain and, ultimately, to the closest A-Topic in the discourse. Interveners in the form of any other overt φ-feature-equipped element, left-dislocated between the NS and C, break the chain and preclude subject drop.

According to Luraghi & Pinelli (2015), in statistical terms, the loss of NSs in Russian evolved simultaneously in root and embedded contexts until approximately the 16th century.[26] Thereafter, NSs in embedded finite clauses experienced a sudden drop, while their decline was much slower in root contexts. Notice that the breaking point, the 16th century, was right at the same time in which overt non-emphatic pronouns became the norm for 1st and 2nd person. Claudi (2014) confirms the same observation for 3rd person *onъ*, which was reanalyzed as a personal pronoun in subordinate finite clauses (complement and obviation structures) much faster than in root clauses.

These statistical data are consistent with the account presented here. Replacement of subject drop by null Topics (root clauses) in successive generations of speakers had to be progressive, i.e. the decay of NSs was very slow in root contexts. Why? On the one hand, pronominal Topics just need the presence of common and widespread informational features at C to be licensed and, on the other, they do not need to be realized always as overt. In partial NS languages, due to the lack of a D-feature on T, speakers always have the option to use the overt realization of an overt subject, whereas the null variant is also available, depending on the specific pragmatic context in which the Topic occurs.

---

26 An anonymous reviewer points out that V2 languages behaved diachronically in a different way as compared to Russian, since the former displayed more NSs in root contexts than in embedded clauses (Cognola & Walkden 2019). This data follows on straightforwardly from the role of V raising to a high position; as argued in this chapter, a high V can help in analyzing T as a D-linker or a mediator between C and the NS in referential index transmission (in a similar way to consistent NS languages). However, Russian was never a V2 language and, once V-to-T movement had been lost, the replacement of NSs by overt subjects was at first similar in embedded and root clauses, as V stayed within vP and could not possibly be interpreted as a D-linker any more.

As for embedded NS, learners followed the model of the second type of referential embedded "subject gap" occasionally found in the language, namely, PRO. Referential PRO, which was present in some volitive structures (see (19)), was bound by a close coreferential c-commanding antecedent in the matrix clause. Following the loss of the unique mechanism of licensing NSs in Grammar 1, "control" was virtually the only available mechanism to interpret embedded referential NSs in Grammar 2. But this implies that embedded NSs became severely restricted by a series of new requirements (coreference, the presence of a close c-commanding antecedent, no split antecedents, etc) that root null Topics are not subject to.[27]

As was noted earlier, the inflection point for the change was the 16th century, when the replacement of Grammar 1 by Grammar 2 in 1st and 2nd person was complete. The imposition of control to non-finite clauses had a further consequence. In particular, it led to the loss of overt embedded dative subjects, a loss which took place in the transition between the 16th and 17th century (Borkovskij 1949, 1978; Ivanov 1990; Madariaga 2011). All infinitive clauses in complement position with coreferential subjects were reanalyzed as instances of control. Non-coreferential infinitive subjects, available in Old Russian in declarative and desiderative infinitive clauses (see (17)), could not be reinterpreted in terms of control, because of the lack of a c-commanding antecedent in the matrix clause. In this case, the whole infinitive structure was replaced by an alternative finite construction, i.e. finite complement clauses with declarative verbs (see (18)), and finite obviation structures with volitive verbs (Madariaga 2015).

These remarks are confirmed by the texts. From the 16th century on, we observe a sudden increase of overt pronominal subjects in finite embedded clauses, especially in the newly created obviation structures with *čtoby*, and declarative clauses with *čto* (Claudi 2014; Luraghi & Pinelli 2015). This follows in a straightforward way from the fact that in Grammar 2 non-coreferent subjects in embedded clauses do not have a close c-commanding antecedent, so they had to be overtly realized in a mandatory way.

The only instances of silent subjects in finite embedded clauses that "survived" in Grammar 2 were those residual coreferential embedded NSs that had been reanalyzed as "finite control", which appeared in a very restricted syntactic

---

27 Control as the mechanism licensing embedded NSs, spread in the same way to finite and infinitive constructions with coreferent subjects (cf. Madariaga 2015). I do not intend to identify embedded NSs in finite and nonfinite contexts completely. I just contend that the finite variants are a historical *extension* of nonfinite NSs = PRO, resulting in a pool of shared properties between both. Nonetheless, as it is stated in section 2.2, I acknowledge that finite and nonfinite clauses are irreducible to a single phenomenon.

environment, i.e. were c-commanded by a local antecedent. In contrast, all the other types of embedded subjects that were newly created in the language, such as (optionally) overt coreferential G-Topics and (always) overt non-coreferential pronouns, were much less restricted from the point of view of their potential suitable environments and, thus, experienced a sharper increase in the number of occurrences, with the advent of the new system of referential NSs.

5 Conclusion

In this chapter, I reviewed the licensing conditions of referential NSs in Modern Russian root and embedded clauses in Modern Russian (a partial NS language) and Old Russian (a consistent NS language), and explained the change between the two stages.

Our Grammar 2 is represented by Modern Russian. Here, we can distinguish at least three mechanisms of licensing referential NSs. In root contexts, NSs match their unvalued φ-features with T, but still need to get a referential index in order to be correctly interpreted (given that T lacks an unvalued D-feature). This referential index can be transmitted by D-linking in one of the following configurations: (i) G-topical subjects, whenever the corresponding logophoric or situational/contextual features are accessible at C (root or embedded contexts), (ii) root NSs in a topic chain, licensed by a null operator in the C-layer which shares its features, including the referential index, with every link of the chain, the antecedent, and the NS below, and (iii) embedded coreferential NSs, bound by a c-commanding antecedent, in similar conditions in finite and non-finite contexts.

In Old Russian, our Grammar 1, the type of Topic or the degree of embedding (root or subordinate) was not relevant to license NSs and every NS could be dropped given that its reference was recoverable from the context, i.e. that it was bound by a Topic that was able to anchor the reference of the NS. On the other hand, embedded subjects displayed no control, even in non-finite clauses; the only control-like constructions occurred in infinitive clauses with coreferent subjects that performed as complements to some volitive verbs. In Old Russian, T was endowed with an unvalued D-feature, and it was V-T, located high in the structure, which "mediated" in referential index transmission from a null Topic/operator at CP to the φP (NS), when T valued its own D-feature.

The change from Grammar 1 into Grammar 2 started with the loss of V-to-T movement, which left the T position empty, canceling the ability of T to mediate as a referential index transmitter from CP to the referential NS (i.e. T was reanalyzed as having no unvalued D-feature). Under regular circumstances, pronominal

subjects had to be realized as overt. However, residual instances of the old system of subject drop did not disappear, but were reanalyzed by learners of Grammar 2. The reanalysis was performed in different ways depending on the specific configurations, as far as learners were able to interpret the reference of the subject gap. Thus, three mechanisms were put into use: (i) null G-Topics, when the adequate features at CP (logophoric or situational/contextual features) were available; (ii) topic chains, in case the NS could be interpreted as a coreferential successive copy of an overt A-Topic; (iii) anaphoric embedded NSs, i.e. those bound by an antecedent in the matrix clause.

The last one of these was an already existing mechanism in the language in some volitive infinitive structures with coreferential subjects. The reanalysis of residual coreferential embedded NSs as "controlled" was performed in non-finite and finite clauses in a parallel way, only that in the case of non-finite clauses, the immediate consequence was the loss of overt infinitive dative subjects (which fell outside the requirements of control), while in finite control the newly reanalyzed controlled NSs could alternate with the old and the newly created (non-controlled) overt pronominal subjects.

References

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parameterizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16 (3). 491–539.

Alonso-Ovalle, Luis, Susana Fernández-Solera, Lyn Frazier & Charles Clifton. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Rivista di Linguistica* 14 (2). 1–19.

Bailyn, John. 2012. *The Syntax of Russian*. Cambridge: Cambridge University Press.

Barbosa, Pilar. 2019. *Pro* as a minimal nP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50 (3). 487–526.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge Handbook of Historical Syntax*, 134–162. Cambridge: Cambridge University Press.

Bobaljik, Jonathan. 2002. A-chains at the PF-interface: Copies and covert movement. *Natural Language and Linguistic Theory* 20 (2). 167–267.

Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.

Borkovskij, V. I. 1949. *Sintaksis drevnerusskix gramot. Prostoe predloženie* [Syntax of Old Russian letters. The simple sentence]. L'vov: Izdatel'stvo L'vovskogo gosudarstvennogo universiteta.

Borkovskij, V. I. 1978. *Istoričeskaja grammatika russkogo jazyka: sintaksis – prostoe predloženie* [Historical grammar of Russian. Syntax – the simple sentence]. Moscow: Nauka.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36 (1). 1–22.

Claudi, Tommaso. 2014. *The Status of Subject Pronouns in Old Russian. A Diachronic Analysis*. Pavia: University of Pavia MA thesis.

Cognola, Federica & George Walkden. 2019. *Pro*-drop in interrogatives and declaratives: a parallel study of Old High German and Old Italian. *Linguistik Online* 100 (7). 95–140.

Eckhoff, Hanne & Roland Meyer. 2011. Conditions on null subjects in Old Church Slavonic, a contrastive view. Paper presented at the 20 ICHL, Osaka.

Ferreira, Marcelo. 2009. Null subjects and finite control in Brazilian Portuguese. In Nunes, Jairo (ed.), *Minimalist Essays on Brazilian Portuguese Syntax*, 17–49. Amsterdam/Philadelphia: John Benjamins.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential pro. *Natural Language and Linguistic Theory* 25 (4). 691–734.

Frascarelli, Mara. 2018. The interpretation of pro in consistent and partial null subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar. A Synchronic and Diachronic Perspective*, 211–239. Oxford: Oxford University Press.

Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language and Linguistic Theory* 31(1). 91–136.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36 (4). 533–564.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63 (1): 59–97.

Hornstein, Norbert. 1999. On control. *Linguistic Inquiry* 30 (1). 69–96.

Ivanov, V. V. 1990. *Istoričeskaja grammatika russkogo jazyka* [Historical grammar of Russian]. Moskva: Prosveščenie.

Jiménez-Fernández, Angel. 2016. When discourse meets null subjects. *Borealis: An International Journal of Hispanic Linguistics* 5 (2). 173–189.

Jung, Hakyung. 2018. Null subjects and person in Old North Russian. In Jasmina Grkovic-Major, Björn Hansen & Barbara Sonnenhauser (eds), *Diachronic Slavonic Syntax: The Interplay between Internal Development, Language Contact, and Metalinguistic Factors*, 95–123. Berlin: Mouton de Gruyter.

Karskij, Je. F. (ed.). 2001 [1926–28]), *Lavrent'evskaja letopis' – Suzdal'skaja letopis'. Polnoe sobranie russkix letopisej – tom I* [Laurentian chronicle – Suzdal chronicle. Collection of Russian chronicles – volume 1]. Moscow.

Kibrik, Andrej A. 2013. Peculiarities and origins of the Russian referential system. In Dik Bakker & Martin Haspelmath (eds.), *Languages across Boundaries: Studies in Memory of Anna Siewierska*, 227–262. Berlin: Mouton de Gruyter.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Working Papers of SFB 632. Interdisciplinary Studies on Information Structure (ISIS) 6*, 13–56. Potsdam: Universitätsverlag Potsdam.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22 (4). 811–877.

Landau, Idan. 2006. Severing the distribution of PRO from case. *Syntax* 9 (2). 153–170.

Landau, Idan. 2015. *A Two-tiered Theory of Control*. Cambridge, MA: MIT Press.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.

Livitz, Inna. 2014. *Deriving silence through internal reference: focus on pronouns*. PhD dissertation. New York University, New York.

Luraghi, Silvia & Erica Pinelli. 2015. The loss of referential null subjects in Russian: Can subordinate clauses can tell us? Paper presented at the conference on Slavic Corpus Linguistics, Tromsø.

Madariaga, Nerea. 2011. Infinitive clauses and dative subjects in Russian. *Russian Linguistics* 35 (3). 301–29.

Madariaga, Nerea. 2015. The decline of non-finiteness as a mechanism for embedding in East Slavic. *Journal of Historical Linguistics* 5 (1). 139–174.

Madariaga, Nerea. 2018. Diachronic change and the nature of pronominal null subjects: the case of Russian. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar. A Synchronic and Diachronic Perspective*, 171–98. Oxford: Oxford University Press.

Meyer, Roland. 2011. *The History of Null Subjects in East Slavonic. A Corpus-based Diachronic Investigation*. Habilitation thesis. University of Regensburg, Regensburg.

Migdalski, Krzysztof. 2013. Diachronic source of two cliticization patterns in Slavic. In Christine Meklenborg-Salvesen & Hans Petter Helland (eds.), *Challenging Clitics*, 135–158. Amsterdam: John Benjamins.

Modesto, Marcello. 2011. Finite control: Where movement goes wrong in Brazilian Portuguese. *Journal of Portuguese Linguistics* 10 (2). 3–30.

Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence Topics. *Philosophica* 27 (1). 53–94.

Roberts, Ian. 2009. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 58–87. Cambridge: Cambridge University Press.

Roberts, Ian. 2010. *Agreement and head movement*. Cambridge, MA: MIT Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Michelle Sheehan & Ian Roberts, *Parametric Variation. Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Rodrigues, Cilene. 2004. *Impoverished Morphology and A-movement Out of Case-domains*. PhD dissertation. University of Maryland, MD.

Shushurin, Philip. 2017. Null pronouns in Russian embedded clauses. In Pritty Patel-Grosz, Patrick Georg Grosz & Sarah Zobel (eds.), *Pronouns in Embedded Contexts at the Syntax-Semantics Interface*, 145–169. Berlin: Springer.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument *pro*-drop. *Linguistic Inquiry* 42 (2). 267–304.

Tsedryk, Egor. 2012. Finite control in Russian as movement: Subject extraction, feature deficiency and parametric variation. Ms., Saint Mary's University.

Tsedryk, Egor. 2015. Deriving null pronouns: A unified analysis of subject drop in Russian. In Małgorzata Szajbel-Keck, Roslyn Burns & Darya Kavitskaya (eds.), *Formal Approaches to Slavic Linguistics 23: The First Berkeley Meeting*, 342–361. Ann Arbor: Michigan Slavic Publications.

Tsedryk, Egor. 2022. Null subjects and their overt counterparts in East Slavic root clauses: referential and non-referential readings. In Gréte Dalmi, Egor Tsedryk & Piotr Cegłowski (eds.),

Null Subjects in Slavic and Finno-Ugric. Licensing, Structure, and Typology, 35–74. Berlin: Mouton de Gruyter.

Vainikka, Anne Marjatta & Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17 (3). 613–671.

Zaliznjak, A. A. 2004. *Drevnenovgorodskij dialect* [Old Novgorod Dialect]. Moskva: Jazyki slavjanskoj kul'tury.

Zaliznjak, A. A. 2008. *Drevnerusskie enklitiki* [Old Russian Clitics]. Moskva: Jazyki slavjanskoj kul'tury.

Marta Ruda

4 Interpreting null subjects in Polish: Against left-peripheral linking

1 Introduction

A number of recent discussions related to the interpretation of null subjects in consistent Null Subject Languages (NSLs) point to the observation that these subjects are interpreted with reference to a topic, understood to be introduced into the syntactic derivation in a designated position in the left periphery of the clause (Frascarelli 2007, 2018; Frascarelli & Jiménez-Fernández 2019; Holmberg & al. 2009; Holmberg 2010; Sigurðsson 2011, 2014, 2019; Jiménez-Fernández 2016, among others).[1] Within this line of research, the position relevant for licensing and interpreting null subjects is the syntactic Aboutness-Shift Topic position, the highest in the hierarchy of distinct topic positions represented in the clausal spine in the cartographic approach to clausal structure (Rizzi 1997). The About-ness-Shift Topic c-commands Contrastive Topic, which c-commands About-ness-Given Topic, which, in turn, c-commands Familiar-Given Topic.[2] Technical, syntactic machinery is thus developed to account for the association of the null subject with the Aboutness-Shift Topic, including the far-reaching proposal of imposing the Topic Criterion, provided here in (1). The Topic Criterion is meant to hold in all consistent NSLs, taking the form of a macroparameter, that is, a parameter distinguishing between non-*pro*-drop languages and *pro*-drop languages (of the consistent and partial type; see Frascarelli 2018).

(1) *Topic Criterion* (revised; Frascarelli & Jiménez-Fernández 2019: 32)
   a. A third-person argument *pro* matches the [aboutness] [ref] features that are encoded in the A-Topic projection (ShiftP) in the C-domain.
   b. A-Topic chains can only be started in root domains (the Interface Root Restriction).
   c. A-Topic chains across sentences require local links, which can be overt or silent.
   d. A-Topics can be silent.

In what follows, I employ data from Polish to show that the picture of licensing and interpretating null subjects in consistent NSLs is much more complex than what can be derived from the Romance and Germanic data. I start the discussion in section 2 by offering some general remarks about Polish as a consistent NSL. In section 3, I move on to considering the Topic Criterion (especially the relation between null subjects and the Aboutness-Shift Topic) from the point of view of the relevant data from Polish. I conclude that an approach along these lines is inadequate to account for the data and, what follows, the Topic Criterion is not a viable candidate for a macroparameter. In this connection, I also consider the role of agreement and contextual and world knowledge factors for the interpretation of null subjects. In section 4, I look more closely at the broad interpretive properties of null subjects in Polish. Section 5 concludes the chapter with some general theoretical remarks concerning the Topic Criterion.

2 Polish as a consistent NSL

Polish is a language with the basic SVO, although in general highly flexible, word order (see Willim 1989 for a discussion; and Haider (to appear), for a different take on word order phenomena in languages such as Polish). It is a consistent NSL and has been classified as such in some previous literature (see Barbosa 2019: 520, *fn* 16; Roberts 2019: 250; and the references therein). Null subjects are available with all person/number/gender combinations and all tense/aspect/mood combinations. In the present tense, verbal morphology reflects the person and number of the subject with a distinct fusional morpheme for all person/number combinations, as (2) illustrates.

(2) a. *Czyta-m.*
         read-1SG
         'I read/am reading.'
     b. *Czyta-sz.*
         read-2SG
         'You read/are reading.'
     c. *Czyta.*
         read.3SG
         'She/he/it reads/is reading.'
     d. *Czyta-my.*
         read-1PL
         'We read/are reading.'

    e.   *Czyta-cie.*
        read-2PL
        'You read/are reading.'

    f.   *Czyta-ją.*
        read-3PL
        'They read/are reading.'

In structures where the lexical verb appears in the so-called *l*-participle form (e.g. in the past tense) and the *n/t*-participle form (the passive voice), also gender is morphologically expressed, as in (3).[3]

(3) a. *Czytał-a-m.*
        read-SG.F-1SG
        'I (have) read.'

    b. *Czytał-e-m.*
        read-SG.M-1SG
        'I (have) read.'

    c. *Będzie*   *czytan-a.*
        will.be.3SG   read-SG.F
        'She (it) will be read.'

    d. *Będzie*   *czytan-y.*
        will.be.3SG   read-SG.M
        'He (it) will be read.'

As can be expected, Polish does not use overt expletive subjects, as (4) illustrates for a *weather*-predicate (a context for a quasi-argumental expletive), (5) for a raising predicate and (6) for the existential context.[4]

---

3 Polish is a language with grammatical gender, which is why the feminine form in (3c) could be used, for example, with the intended subject interpreted as *książka* 'book.F', whereas the masculine form in (3d) could be used with *wiersz* 'poem.M'. For simplicity's sake, *l*-participles and *n/t*-participles are not marked explicitly in the glosses.

4 Franks (1995: 322) provides some examples which he interprets as marginally making an expletive *to* 'this' subject possible. However, this point requires further investigation, in order to establish whether *to* in these cases is indeed a subject expletive. One alternative might be a demonstrative introducing the finite clause, with extraposition separating these two; another one is a predication marker, depending on the particular examples. I thank an anonymous reviewer for drawing my attention to Franks's data.

(4) (*To/ *ono) *pada.*
     this   it   rains
     'It is raining.'

(5) (*To/ *ono) *wydaje się, że   będzie padać.*
     this   it   seems SE   that will   rain
     'It seems that it will rain.'

(6) (*To/ *ono/ *te/ *one) *pojawiły   się problemy.*
     this   it   these they   appeared SE   problems
     'Problems have appeared.'

In line with Holmberg's (2010: 91) discussion, Polish, as a consistent NSL, requires the embedded subject of a finite clause to be null, if co-reference with the matrix subject is intended, as (7) shows.[5]

(7) a. *Jan*ᵢ *powiedział, że   (*onᵢ) chce   kupić samochód.*
         Jan said.SG.M   that   he   wants buy   car
         'Jan said that he wanted to buy a car.'
     b. *Jan*ᵢ *nic   nie powiedział, ale Paweł*ⱼ *powiedział, że*
         Jan nothing not said.SG.M   but Paweł said.SG.M   that
         ___ᵢ/ⱼ *chce   kupić samochód.*
         (he) wants buy   car
         'Jan didn't say anything, but Paweł said that he wanted to buy a car.'

However, as I show in what follows, the interpretive options available to null subjects in Polish are broader than what has been reported for some other consistent NSLs. Before I move on to this issue though, a note on the generic reading is in order, this being an important component of Holmberg's (2005, 2010) approach to the classification and analysis of NSLs.

In particular, Holmberg observes that consistent NSLs do not make the generic (speaker-inclusive) reading available with a third person singular null subject, an aspect of the analysis which Roberts (2019) refers to as Holmberg's

---

5 This holds of non-focused pronominal subjects. When the subject of the embedded clause is focused, it can co-refer with the matrix subject despite being overt, as in (i).

(i) *Jan*ᵢ *powiedział, że   tylko onᵢ chce   kupić samochód.*
     Jan said   that   only he wants buy   car
     'Jan said that only he wanted to buy a car.'

Other Generalisation. This is true for Polish with respect to examples such as (8), where in (8a) singular agreement on the verb is compatible only with the definite, third person singular reference. Plural agreement is compatible both with definite third person plural reference and with arbitrary (speaker-exclusive) reference, as (8b) shows. To achieve generic reference, the SE marker *się* and third person singular ragreement is used, as in (8c).

(8) a. *Tak przygotowuje ten deser.*
          thus prepare.3SG this dessert
          'This is how she/he/*one prepares this dessert.'
     b. *Tak przygotowują ten deser.*
          thus prepare.3PL this dessert
          'This is how they/people prepare this dessert.'
     c. *Tak przygotowuje się ten deser.*
          thus prepare.3SG SE this dessert
          'This is how one prepares this dessert.'

However, as (9) shows, the generic reading of a third person singular subject is also available in Polish if the subject is embedded in a previously established generic context (see Dalmi 2017 for a discussion of a similar pattern in Hungarian).

(9) a. *Jak człowiek nie śpi wystarczająco, to __*
          if man not sleeps enough then (man)
          traci odporność
          loses immunity
          'If one doesn't sleep enough, then one loses immunity.'
     b. *Człowiek nigdy nie wie, kiedy __ zostanie*
          man never not knows when (man) will.become.3SG
          w tym sklepie oszukany.
          in this shop cheated
          'One never knows when one will be cheated in this shop.'

What is more, apart from the second person singular generic reading illustrated in (10), Polish also has at its disposal a structure built around defective modal predicates such as *można* 'be allowed to/may', *należy* 'should', *warto* 'worth', and *trzeba* 'ought to', which do not inflect for the subject [φ]. A structure of this type can be used with the generic interpretation and the availability of anaphoric binding, as in (11c), testifies to the presence of a syntactically represented subject in the structure (anaphors in Polish are strictly subject oriented).

(10) *Jak* __ *chcesz chodzić zimą po górach,* __
if (you) want.2SG walk.INF winter.IN in mountains (you)
musisz mieć swoje raki.
have.to.2SG have.INF self's crampons
'If you want to hike in the mountains in winter, you have to have your own crampons.'

(11) a. *Tu* __ *nie można palić.*
here (one) not may smoke.INF
'One may not smoke here.'
 b. *W Anglii* __ *należy mówić po angielsku.*
in England (one) should speak.INF in English
'One should speak English in England.'
 c. __ *Trzeba pomagać swoim przyjaciołom.*
(one) ought.to help.INF self's friends
'One ought to help one's friends.'

The morphological marking on the modifier of the anaphor in (12) shows that the subject can be third person singular masculine (masculine is the semantically unmarked member of the masculine/feminine pair for human nouns).

(12) __ *Trzeba szanować siebie samego.*
(one) ought.to respect.INF self alone.SG.M.ACC
'One ought to respect oneself.'

The forms of these predicates used in structures with past temporal reference bear (default) 3SG.N agreement marking. Some of them require an auxiliary, as in (13a, c), and some inflect on their own, as in (13b).

(13) a. *Tu* __ *nie można było palić.*
here (one) not may was.3SG.N smoke.INF
'One could not smoke here.'
 b. *W Anglii* __ *należało mówić po angielsku.*
in England (one) should.3SG.N speak.INF in English
'One needed to speak English in England.'
 c. __ *Trzeba było pomagać swoim przyjaciołom.*
(one) necessary was.3SG.N help.INF self's friends
'One had to help one's friends.'

Thus, what is relevant for Holmberg's Other Generalization is only third person singular non-default morphology in contexts such as (8) (though not as in (9)).[6] While these facts are of non-trivial theoretical interest, the semantic and syntactic aspects of analysing generic, arbitrary and impersonal structures are highly complex and this issue is not pertinent to the main point of this contribution. This is why I will not discuss it further here, but merely refer the interested reader to some relevant discussions of Polish in Kibort (2001, 2008), Lavine (2005, 2017), Krzek (2010, 2013, 2017) and Ruda (2014).

3 Null subjects in Polish and Aboutness-Shift Topics

Based on data from Italian, Greek, Hebrew and Chinese, Samek-Lodovici (1996) proposes the generalisation that the antecedents of null subjects must be topics. Working with Italian data, Frascarelli (2007) identifies the topic type as the Aboutness-Shift Topic. While she discusses some naturalistic data which fit this pattern, the main piece of evidence offered for the approach on which null subjects have to be interpreted via linking with the Aboutness-Shift Topic is the unavailability of a null subject in the presence of an overt, non-coreferential Aboutness-Shift Topic in the sentence in Italian. As Aboutness-Shift Topic is a unique, non-iterative position in the sentence, the subject is predicted to always co-refer with it. The Polish data in (14) thus instantly reveal that this hypothesis cannot account for the entire spectrum of null subjects in all consistent NSLs, the null

---

6 While Krzek (2017) suggests that the generic inclusive reading is available in the so-called impersonal *-no/to* construction in Polish, usually associated with the arbitrary (speaker- and addressee-exclusive) reading, the relevant examples cited by her actually involve what Holmberg & Phimsawat (2017) refer to as quasi-inclusive reading, either the speaker or the addressee not being necessarily included in the interpretation of the null subject. In any case, the *-no/to* construction can be argued to involve default agreement morphology (Ruda 2014), in parallel to the examples featuring defective predicators discussed in the main text above, and a plural rather than singular subject. I thank an anonymous reviewer for drawing my attention to Krzek (2017).

subject in (14) referring to the coach, despite the presence of the Aboutness-Shift Topic introducing Anna.[7]

(14) *Trener*$_i$ *podsumował dziś postępy zawodniczek. Co do Anny*$_j$,
      coach   summed_up today progress players.GEN what to Anna
      __$_i$ *był niezadowolony z jej*$_j$ *wyników.*
      (he) was dissatisfied.SG.M from her results
      'The coach has summed up the players' progress today. As for Anna, he was dissatisfied with her results.'

While this single example is enough to show that the Topic Criterion and its elevation to the position of a UG parameter are not steps in the right direction (see also Sigurðsson 2019), in what follows I present a number of other contexts illustrating the broad spectrum of interpretive options available to null subjects in Polish, supporting this point even further. A crucial aspect of the Topic Criterion is that Aboutness-Shift Topics can be silent. In this context, it is worthwhile to compare some relevant Italian data with Polish. In particular, to substantiate the claim that Aboutness-Shift Topics can be null, Frascarelli (2007: 710) provides the data in (15), intended to show that when an overt Aboutness-Shift Topic is present in the structure, the null subject has to be bound by it, an element of the proposal which follows from the hypothesis that Aboutness-Shift Topics are not iterative (i.e. the C-domain can include only one).

(15) a. *Conduco insieme al capo questa rubrica*$_k$ . . .
          'I conduct with my boss this program.'
      b. . . . *la prima parte*$_j$, *fortunatamente,* __$_j$ *è più facile.* __$_{*k/j}$ *è scritta.*
          'the first part, luckily, is easier. It is written.'
      c. . . . *la prima parte*$_j$, *l'*$_j$ *ho già preparata.* __$_{*k/j}$ *è più facile.* __$_{*k/j}$ *é scritta.*
          'I have already prepared the first part. It is easier. It is written.'

---

[7] In a number of examples throughout the paper, including (14), an overt pronominal subject can also be used in the post-verbal position and may be preferred by some speakers, especially in the spoken register. The observation that, in some contexts, Polish makes it possible for overt, mostly post-verbal, pronominal subjects to be used instead of null subjects with no difference in meaning is noted in Pisarkowa (1969), Nillson (1982) and McShane (1999, 2009). In general, the spoken variety of Polish shows a greater frequency of use of overt subject pronouns with apparent lack of a meaning difference, a situation opposite to Russian (McShane 2009). The contexts which make both options possible require much more further research, both descriptive and theoretical, and cannot be addressed in greater detail here. What is important for the present purpose is that structures such as (14) are generated by the grammar of Polish, at least for the speakers whom I have been able to consult so far.

Neither in (15b), where the topic is the subject, nor in (15c), where it is the object resumed by a clitic, can the null subject be interpreted as coreferential with *questa rubrica* 'this program', a Familiar Topic at this point. Furthermore, Frascarelli (2007: 711) reports that a mismatch in φ-features between the new Aboutness-Shift Topic and the null subject, as evidenced by verbal morphology, leads to a situation in which the null subject cannot be interpreted either as coreferential with the Familiar Topic or with the Aboutness-Shift Topic, as in (16).

(16) *<questa rubrica> conduco insieme al capo questa rubrica$_k$. Le prime due parti$_j$, fortunatamente, __$_j$ sono finite. __$_{*k/*j}$ é scritta.*
'I conduct with my boss this program. Luckily, its first two parts are finished. #It is written.'

The situation in Polish is different. While the most natural interpretation of (17a) is the one on which the null subject is coreferential with *segment*, both *segment* and *program* being of masculine gender in Polish and the second clause being most naturally interpreted as an explanation for why the first segment is easier, in (17b), where the feminine noun *część* 'part' is used instead of *segment*, the co-reference with the Familiar Topic, ignoring the Aboutness-Shift Topic, is possible, especially if embedded in a more favourable context. As (17c) shows, adding this context to (17a) removes to some extent the interpretation of the second sentence as providing a reason for the easiness of the first segment and therefore makes co-reference with *ten program* 'this program' available. The same holds of the mismatch in number shown in (17d).[8]

(17) *Prowadzę ze   swoim szefem ten   program$_i$.*
run.1SG   with self's boss   this program
'I'm running this program with my boss.'
a. *Jego pierwszy segment$_j$ jest na szczęście łatwiejszy.*
his   first   segment.M   is   on fortune   easier.SG.M
__$_j$ *jest spisany.*
(he) is   written.SG.M
'Fortunately, its first segment is easier. It is written.'

---

8 A reviewer reports 'mixed feelings' about (17). I present it here to be as close to Frascarelli's data as possible. I offer more natural contexts in what follows.

b. *Jego pierwsza część*$_k$ *jest na szczęście łatwiejsza.* __$_i$
his first part.NOM.F is on fortune easier.SG.F (he)
jest (*też ogólnie wcześniej*) *spisany.*
is also generally beforehand written.SG.M
'Fortunately, its first part is easier. (Generally) It is (also pre-)written.'

c. *Jego pierwszy segment*$_j$ *jest na szczęście łatwiejszy.*
his first segment.NOM.M is on fortune easier.SG.M
__$_{?i/j}$ *jest* (*też ogólnie wcześniej*) *spisany.*
(he) is also generally beforehand written.SG.M
'Fortunately, its first segment is easier. (Generally) It is (also pre-)written.'

d. *Jego dwie pierwsze części*$_l$ *są na szczęście łatwiejsze.*
his two first parts are on fortune easier.PL
__$_i$ *jest* (*też ogólnie wcześniej*) *spisany.*
(he) is also generally beforehand written.SG.M
'Fortunately, its first two segments are easier. (Generally) It is (also pre-)written.'

For the sake of completeness, (18) provides a context for the pattern in which the Familiar Topic is plural and the Aboutness-Shift Topic singular.

(18) *Prowadzę ze swoim szefem te programy*$_i$. *Ich pierwsza*
run.1SG with self's boss these programs their first
część$_j$ *jest na szczęście łatwiejsza.* __$_i$
part.F is on fortune easier.SG.F (they)
są (*też ogólnie wcześniej*) *spisane.*
are also generally beforehand written.PL
'I'm running these programs with my boss. Fortunately, their first part is easier. (Generally) They are (also pre-)written.'

The same is observed when the Aboutness-Shift Topic precedes the sentence, as in (20), based on Frascarelli's (2007: 711) (19), which is reported to have only one available interpretation. On the other hand, the null subject in the final sentence in (20a) is ambiguous in this context and no issues arise when reference to the program is intended. The interpretation can be disambiguated by the mismatch in number with the introduced Aboutness-Shift Topic, as (20b) illustrates.

(19) *L'intervista*$_j$ *la*$_j$ *sto preparando perché conduco insieme al capo questa rubrica*$_k$. __$_{j/*k}$ *scritta.*
'As for this interview, I'm preparing it because I conduct with my boss this program. It is written.'

(20) a. *Jeżeli chodzi o   ten   wywiad*_i, *przygotowuję go*_i, *bo*
          if   goes about this interview prepare.1SG him because
          *prowadzę ze   swoim szefem ten   program*_j. ____{i/j} *jest* (*już*
          run.1SG with self's boss   this   program   (he) is   already
          ogólnie)   *spisany.*
          generally   written.SG.M
          'As for this interview, I'm preparing it, because I'm running this
          program with my boss. (Generally) It is (already) written up.'
      b. *Jeżeli chodzi o   te   wywiady*_i, *przygotowuję je*_i,
          if   goes about these interviews prepare.1SG them
          *bo   prowadzę ze   swoim szefem ten   program*_j. ____j
          because run.1SG with self's boss   this   program   (he)
          jest (*już   ogólnie*)   *spisany.*
          is   already generally written.SG.M
          'As for these interviews, I'm preparing them, because I'm running this
          program with my boss. (Generally) It is (already) written up.'

Another context employed in the literature to show that the third person null
subject needs to be interpreted with reference to the Aboutness-Shift Topic
involves co-reference with the noun phrase introduced in a *by*-phrase inside
a passive structure, as in (21), discussed in Holmberg (2010: 96), following
Samek-Lodovici (1996) and Frascarelli (2007).

(21) a. *Questa mattina, la   mostra   è stata visitata di   Gianni.*
          this   morning the   exhibition was   visited by   Gianni
          *Pìu tardi *___ / egli / lui ha   visitato l'università.*
          later   (he) / he / he PERF visited the.university
          'This morning the exhibition was visited by Gianni. Later he visited
          the university.'
      b. *Questa mattina, Gianni ha   visitato la   mostra.   Pìu tardi*
          this   morning Gianni PERF visited the   exhibition later
          ___ *ha   visitato l'università.*
          (he) PERF visited the.university
          'This morning Gianni visited the exhibition. Later he visited the
          university.'

The equivalent of (21a) is also highly unnatural in Polish (see (22a)), but actu-
ally the first sentence in its own right is rather awkward, even if grammatically

well-formed, and neither option seems perfect.[9] When the context is altered slightly to remedy this issue, a null subject is actually completely acceptable in environments of this type in Polish, as in (22b).

(22) a. *Tego   ranka   wystawa   została   odwiedzona   przez   Jana.*
         this   morning   exhibition was   visited   by   Jan
         Potem   __   poszedł   na   uniwersytet.
         later   (he)   went   on   university
     b. *Pół   roku   temu   Anna   została   zauważona   przez   słynnego*
         half   year   ago   Anna   was   noticed   by   famous
         fotografa   W   zeszłym   tygodniu   __
         photographer   in   last   week   (he)
         zaproponował   jej   nawet   stałą   współpracę.
         offered   her   even   permanent   cooperation
         'Half a year ago Anna was noticed by a famous photographer. Last week he even offered her permanent cooperation.'

Similarly, (23) illustrates that the null subject in Polish can be coreferential with an NP introduced as a possessor (see also Cole 2010 for a discussion of cross-linguistic data, likewise showing that at least in some NSLs the null subject does not need to be coreferential with the topic in similar contexts). In this case, as the embedded clause in the second sentence contains the verb in the present tense, where gender is not marked, the sentence is ambiguous, with the null subject interpreted either as the mom or as Tomek.

(23) __   *Rozmawiałam   z   [mama   Tomka$_j$]$_i$.   __$_i$ żałowała,*
     (I)   talked.1SG.F   with   mom   Tomek.GEN   (she)   regretted.SG.F

     że   __$_{i/j}$ nadal   jest   w   szpitalu.
     that (s/he)   still   is   in   hospital
     'I talked with Tomek's mom. She regretted that s/he is still in hospital.'

---

9 If an overt pronominal subject is introduced, it is better in the post-verbal than the pre-verbal position.

(i) *Tego   ranka   wystawa   została   odwiedzona   przez   Jana.   Potem*
     this   morning   exhibition became   visited   by   Jan   later
     (???*on*) *poszedł   (?on)   na   uniwersytet.*
     he   went   he   on   university
     'This morning the exhibition was visited by Jan. Later he went to the university.'

Furthermore, the null subject in Polish can have the subject of a thetic sentence as an antecedent, even if an overt Aboutness-Shift Topic is present.

(24) A: *Co   się   dzieje?*
       what   SE   happens
       'What is happening?'

   B: *Anna$_i$ jest dziś   strasznie rozkojarzona.*
      Anna   is   today terribly   distracted
      'Anna is terribly distracted today.'

   A: *Ale   jeżeli chodzi o   nasze spotkanie, to   o*
      but   if   goes   about   our   meeting   then about
      nim   __$_i$ nie   zapomniała?
      him   (she) not   forgot.SG.F
      'But as for our meeting, she hasn't forgotten about it?'

In addition, as Frascarelli (2018) points out, the complement of a factive verb, as an element introducing presupposed information, lacks the Aboutness-Shift Topic position by definition. If the null subject is identified with the subject of a higher clause within the complement of a factive verb, the approach under discussion here needs to assume that this subject is an Aboutness-G(iven) Topic, linked to an Aboutness-Shift Topic located in the matrix clause. This is illustrated here in (25b) for the Italian example in (25a) from Frascarelli (2018: 226–227).

(25) a. *Jari si dispiace   che Leo pensa   che __ perderà*
      Jari SE be.sorry.3SG that Leo think.3SG that (he) lose.FUT.3SG
      la   gara.
      the race
      'Jari is sorry that Leo thinks that he will lose the race.'

   b. [$_{ShiftP}$ <*Leo$_z$* > [*Jari$_k$ si dispiace* [$_{CP}$ *che* [$_{FamP}$ *Leo$_z$* [t$_z$ *pensa* [$_{CP}$ *che* [__$_z$ *perderá la gara*]]]]]]]]

In the representation in (25b), the two occurrences of Leo are linked with *pro*, forming a 'Topic-chain'. The first occurrence is a silent Aboutness-Shift Topic, whereas the second occurrence is an overt Abountess G-Topic. However, also in this context the Polish data are problematic to this approach. As (26) shows, the embedded null subject can be coreferential with the subject in the complement of a factive verb, despite the fact that the position of the Aboutness-Shift Topic in the matrix clause is occupied by another, non-coreferential element. In fact, the null subject can be coreferential with either of the higher subjects here.

(26) *Jeżeli chodzi o tę umowę, Tomek*_i *żałuje,* [*że Adam*_j
if goes about this contract Tomek regrets that Adam
myśli, [*że* ____{i/j} *został oszukany przy jej*
thinks that (he) became.SG.M cheated.SG.M by her
podpisywaniu]].
signing
'As for this contract, Tomek regrets that Adam thinks that he was cheated
by signing it.'

The data in (23) above also constitute a problem in this connection. More specifically, as the first null subject in the second sentence is interpreted as Tomek's mom (*mama Tomka*), this would need to be the Aboutness-Shift Topic in this sentence. The availability of the interpretation of the null subject inside the second clause, a complement of a factive verb, as referring to *Tomek* is then predicted to be impossible, contrary to fact. What is more, as Frascarelli (2007) notes, a Topic chain cannot be started by phrases associated with focus. In this light, (27), where the null subject refers to Anna, introduced as a focus, despite the presence of an overt Aboutness-Shift Topic, is likewise unexpected.

(27) A: *Kto dzwonił?*
who called.SG.M
'Who called?'

B: *Anna*_i*. Jeżeli chodzi o naszą piątkową wycieczkę,*
Anna if goes about our Friday trip
____i *mówiła, że też* ____i *chciałaby na*
(she) said.SG.F that also (she) would.like.SG.F on
nią pojechać.
her go
'Anna. As for our Friday trip, she said that she would also like to go on it.'

A similar context in (28)–(29) illustrates further that the understanding of the situation described by the sentence leans the interpretive preference towards one of the possible antecedents or the other, regardless of which one is introduced as the Aboutness-Shift Topic. In particular, both relevant NPs in (28)–(29) (the focused one and the one introduced as Aboutness-Shift Topic) are feminine singular and denote humans, which is why both of them are compatible with the Agent role in a speaking event. The most natural interpretation of the first null subject in (28) is the one on which it is coreferential with the focused phrase *Anna*, in which case

the second null subject needs to be coreferential with topical *Zosia* (otherwise the newly introduced topic would not be linked to any position in the sentence). On the other hand, the interpretational options available to the null subjects in (29) need to be compatible with the interpretation assigned to *jej* 'her', the recipient of the message in the speaking event. If *jej* 'her' is coreferential with *Anna*, the null subject of this clause needs to be coreferential with *Zosia* (a result of the adherence to the conditions on binding, in addition to common sense) and, as a consequence, the most plausible interpretation of the second null subject is the one on which it is coreferential with *Zosia*. On the other hand, on the scenario in which *jej* 'her' is coreferential with *Zosia*, the null subjects are coreferential with *Anna*.

(28) A: *Kto   dzwonił?*
     who   called.SG.M
     'Who called?'

    B: *Anna$_i$. Jeżeli chodzi o   Zosię$_j$, __$_i$ mówiła, że*
     Anna   if   goes   about Zosia (she) said.SG.F that
     już   __$_j$ czuje się lepiej.
     already (she) feels SE better
     'Anna. As for Zosia, she said that she was already feeling better.'

(29) A: *Kto   dzwonił?*
     who   called.SG.M
     'Who called?'

    B₁: *Anna$_i$. Jeżeli   chodzi o   Zosię$_j$, __$_j$ powiedziała jej$_i$,*
     Anna   if   goes   about Zosia (she) told.SG.F   her
     że   już   __$_j$ czuje się   lepiej.
     that already (she) feels SE   better
     'Anna. As for Zosia, she told her that she was already feeling better.'

    B₂: *Anna$_i$. Jeżeli chodzi o   Zosię$_j$, __$_i$ powiedziała jej$_j$,*
     Anna   if   goes   about Zosia (she) told.SG.F   her
     że   już   __$_i$ czuje się   lepiej.
     that already (she) feels SE   better
     'Anna. As for Zosia, she told her that she was already feeling better.'

All these data suggest that the status of the Aboutness-Shift Topic in and of itself has very little to do with the interpretive process of null subjects in Polish. Moreover, even topic change does not require an overt pronoun in Polish, unlike what has been reported for some other NSLs (see Cole 2009). This is illustrated in (30).

(30) Jan_i   czekał.   Anna_j przyjechała. __i   otworzył
     Jan   waited.SG.M Anna   came.SG.F   (he)   opened.SG.M
     jej   *drzwi.*   __j   *weszła*   *bez*   *słowa.*
     her.DAT   door   (she)   entered.SG.F without   word
     'Jan waited. Anna came. He opened the door for her. She entered without
     a word.'

In addition to verbal morphology (see also Section 3.1 below), a major factor influ-
encing the interpretation of null subjects in Polish is the knowledge of the world
(commonsensical understanding of the situation described by the sentence). For
example, while the null subject in (31a) is most plausibly interpreted as coferen-
tial with *Adam*, in (31b) it is interpreted as referring to the little son and in (31c) it
is ambiguous (see Pisarkowa 1969 for an extensive discussion of the interpretive
possibilities observed with null subjects, pronouns, and lexical NPs in contexts
of this type).

(31) a. *Adam_i* *tuli*   [*kilkuletniego* *synka*]_j, *jak*   __i   *wraca*   *z*
        Adam   hugs   few.year.old   son   when   (he)   returns   from
        pracy.
        work
        'Adam hugs his few-year-old son when he returns from work.'
     b. *Adam_i* *tuli*   [*kilkuletniego* *synka*]_j, *jak*   __j   *zaczyna* *płakać.*
        Adam   hugs   few.year.old   son   when   (he)   starts   cry
        'Adam hugs his few-year-old son when he starts crying.'
     c. *Adam_i* *tuli*   [*kilkuletniego* *synka*]_j, *jak*   __i/j *jest*   *w*
        Adam   hugs   few.year.old   son   when   (he)   is   in
        dobrym *chumorze.*
        good   mood
        'Adam hugs his few-year-old son when he is in good mood.'

World knowledge is relevant in (31) because in this case verbal agreement is of no
help in the process of identifying the referent of the null subject. In other cases
though, agreement is a crucial factor, as I show in greater detail in the following
section.

3.1 Agreement morphology

While Frascarelli (2007, 2018) dismisses the idea that agreement morphology plays a role in the interpretation of the null subject, the Polish data clearly show that it can be a crucial factor. The interpretation of the null subject in (32), where, again, it is available despite the presence of a non-coreferential overt Aboutness-Shift Topic, is determined by the feminine or masculine agreement on the verb.

(32) a. *Anna*ᵢ *zwykle* *wygrywa* *z*   *Tomkiem*ⱼ *w* *szachy,* *ale* *jeżeli*
         Anna   usually wins   with   Tomek   in chess,   but if
         chodzi   *o*   *wczorajszą*   *partię, nawet się* __ᵢ   *nie*
         goes   about yesterday.ADJ game   even   SE (she) not
         starała.
         tried.SG.F
         'Anna usually wins with Tomek at chess, but as for yesterday's game, she didn't even try.'
    b. *Anna*ᵢ *zwykle* *wygrywa* *z*   *Tomkiem*ⱼ *w* *szachy,* *ale* *jeżeli*
         Anna   usually wins   with   Tomek   in chess,   but if
         chodzi   *o*   *wczorajszą*   *partię, nawet się* __ⱼ   *nie*
         goes   about yesterday.ADJ game   even   SE (he) not
         starał.
         tried.SG.M
         'Anna usually wins with Tomek at chess, but as for yesterday's game, he didn't even try.'

When agreement is not enough to disambiguate the referent, the null subject is preferably interpreted as coreferential with the preceding subject, even if a different potential antecedent is introduced as an Aboutness-Shift Topic. Co-reference with this topic can most readily be achieved with the use of a demonstrative.

(33) *Anna*ᵢ *zwykle* *zaprasza* *nowych* *pracowników* *na*   *obiad,*
      Anna   usually invites   new   employees   on   dinner
      ale   *jeżeli* *chodzi* *o*   *Zosię*ⱼ,   __ᵢ/??ⱼ/ *ta*₌ᵢ/ⱼ
      but   if   goes   about   Zosia   (she)   this.SG.F
      nie *miała* *na* *to* *ochoty.*
      not had.SG.F on   this desire
      'Anna usually invites new employees to dinner, but as for Zosia, she didn't feel like it.'

This is not to say that ambiguity is not tolerated. It certainly is, as, for example, in (23) and (31c) above.

These data are thus in line with the approach argued for in Cole (2009), who takes agreement morphology to play a role in the process of identifying the referent of the null subject, as opposed to the view that the only relevant factor is the linking of the null pronoun with a topic merged at the left periphery. This is also supported by the data discussed in the following section, where I probe deeper into the nature of the interpretation of null subjects in Polish by considering not only definite, but also indefinite interpretations of the null subject.

4 More on interpretation

Null subjects in Polish are available with n-word antecedents, as in (34a), where the antecedent is *żadnego kota* 'no cat' and the null subject is singular, and in (34b), where the null subject is plural, the intended interpretation concerning cats in general.

(34) a. *Jan nie   nakarmił   żadnego kota, zanim   __*
          Jan   not   fed.SG.M   no   cat   before (he)
          nie   został   umyty.
          not   became.SG.M washed.SG.M
          'Jan hasn't fed any cat before it was washed.'
      b. *Żaden kot nie lubi   być myty.   Tę   niechęć   __*
          no   cat not likes be   washed.SG.M this   reluctance (they)
          mają   pewnie zakodowaną   w   genach.
          have   surely   encoded   in   genes
          'No cat likes being washed. Surely, they have this reluctance encoded in the genes.'

Similarly, a null subject is available in (35), where the antecedent is *nikt* 'nobody', though its interpretation amounts to something like 'the person who would have broken in'.

(35) *Nikt   się tu   na pewno nie włamał. __   musiałby*
      nobody SE   here on   sure   not   broke   (he)   would.have.to.SG.M
      najpierw zhakować system   alarmowy.
      first   hack   system   alarm
      'Nobody broke in here for sure. S/He would have to have hacked the alarm system first.'

The null subject itself cannot be interpreted as an n-word here. This is revealed by examples like (36), where a repetition of the n-word is required, so that we can achieve the interpretation on which nobody passed either exam.

(36) *Nikt*   *nie zdał*   *egzaminu z*   *semantyki we wtorek.*
      nobody   not passed.SG.M   exam   from semantics on Tuesday
      W   *środę*   *też*   #(*nikt*)   *nie zdał*   *egzaminu ze*
      on   Wednesday also   nobody   not passed.SG.M   exam   from
      składni.
      syntax
      'Nobody passed the semantics exam on Tuesday. On Wednesday also nobody passed the syntax exam.'

Similar observations can be made with respect to the quantifier *someone*. For example, the null subject in (37) can only be interpreted as the person who searched through my apartment. Likewise, in (38) the null subject is acceptable only under the interpretation that the same person passed both exams. Otherwise, the quantifier would need to be repeated.

(37) *Wiem,*   *że*   *ktoś*   *przeszukiwał*   *moje mieszkanie,*
      know.1SG that someone searched.SG.M   my   apartment
      bo   __   *zostawił po*   *sobie*   *ślady.*
      because (he)   left.SG.M after   self   traces
      'I know that someone searched my apartment, because s/he left traces.'

(38) *Ktoś*   *zdał*   *egzamin z*   *semantyki we wtorek.*
      somebody passed.SG.M exam   from semantics on Tuesday
      W   *środę*   *też*   (*ktoś*)   *zdał*   *egzamin ze*
      on   Wednesday also   he   passed.SG.M exam   from
      składni.
      syntax
      'Somebody passed the semantics exam on Tuesday. On Wednesday s/he also passed the syntax exam.'

The same holds for free-choice items, which can serve as antecedents when the null subject is not interpreted as quantificational, but not otherwise (see (39) and (40), where the null subject is acceptable only under the interpretation that the same person is assumed to have passed both exams).

(39) *Czy ktokolwiek z nas będzie dziś na zebraniu i czy*
if anyone from us will.be today on meeting and if
__ da radę odpowiednio przedstawić nasz wniosek?
(s/he) manages properly present our proposal
'Will any of us be at today's meeting and will s/he manage to present our proposal properly?'

(40) | *Czy* | *ktokolwiek* | *zdał* | *egzamin* | *z* | *semantyki* |
|---|---|---|---|---|---|
| if | anyone | passed.SG.M | exam | from | semantics |
| *we* | *wtorek* | *i* | *czy* | *w* | *środę* |
| on | Tuesday | and | if | on | Wednesday |
| *(ktokolwiek)* | *zdał* | *egzamin* | *ze* | *składni?* | |
| (he) | passed.SG.M | exam | from | syntax | |

'Did anyone pass the semantics exam on Tuesday and did s/he pass the syntax exam on Wednesday?'

This shows that null subjects in Polish do not have any quantificational force of their own, though they can certainly be indefinite (see (34b) and the discussion around (43)–(47) below).[10] This is similar to what is observed with overt (object) pronouns in the language (see Ruda 2021a,b).

Moving on, an interrogative pronoun can also be an antecedent to a null subject in Polish, as in (41)–(42).

(41) A: *Kto dzwoni do drzwi?*
who rings to door
'Who is ringing at the door?'

B: *Nie wiem, ale __ jest chyba bardzo*
not know.1SG but (he) is apparently very
niecierpliwy.
impatient.SG.M
'I don't know, but s/he is apparently very impatient.'

(42) A: *Co to było?*
what this was
'What was that?'

---

[10] Incidentally, these data also show that null subjects in Polish do not arise as a result of argument ellipsis.

B: *Nie wiem,   ale __ strasznie zawyło.*
   not know.1SG but (it) terribly   howled.SG.N
   'I don't know, but it howled terribly.'

Interestingly, not only definite, but also indefinite interpretations (including non-specific/narrow-scope indefinite readings) are available to the null subject in Polish. The examples in (43)–(46) illustrate this point.

(43) *Anna bardzo   lubi książki,   szczególnie jak   __   są*
     Anna very.much likes books.ACC especially   when (they) are
     pisane   wyrafinowanym językiem.
     written.PL sophisticated   language
     'Anna likes books very much, especially when they are written in sophisti-
     cated style.'

(44) *Anna   lubi piec ciasta drożdżowe. __   zawsze jej*
     Anna   likes bake cakes   yeast   (they)   always her
     wychodzą.
     come.out.3PL
     'Anna likes baking yeast cakes. They always come out right for her.'

(45) *Cały dzisiejszy poranek   __ obsługiwałam   niemiłych   klientów.*
     all today   morning (I) attended.1SG.F unpleasant clients
     całe popołudnie też   __   non-stop przychodzili.
     all afternoon   also (they) non-stop came.PL.M
     'All morning today I attended unpleasant clients. All afternoon they also
     kept coming non-stop.'

(46) A: *Dziś rano   dzwonił   do mnie jakiś ankieter.*
        today morning called.SG.M to me   some   pollster
        'Some pollster called me today in the morning.'
     B: *Do mnie __   nie dzwonił.   Na szczęście, bo*
        to me   (he/none) not called.SG.M on fortune   because
        __ byłam   w strasznym humorze.
        (I) was.1SG.F in terrible   mood
        'He didn't call me./No pollster called me. It is fortunate because I was
        in a terrible mood.'

The availability of both the definite and the indefinite interpretations results in ambiguous contexts such as (47), where there could be two different roe deer or just one.

(47)   *Pod*   *mój*   *dom*   *przyszła*   *dziś*   *sarna.*   *Pod*   *dom*
       under   my   house   came.SG.F   today   roe.deer   under   house
       Anny   *też*   __   *przyszła.*
       Anna.GEN   also   (she)   came.SG.F
       'A roe deer came near my house today. It/one came near Anna's house too.'

Finally, I would like to complete the present empirical discussion by noting that both definite and indefinite null subjects can have extralinguistic antecedents in Polish. In (48), the null subject starts the discourse and can be interpreted due to the shared knowledge of the events preceding one of the interlocutor's entering the office. In (49), just as in (47), both definite and indefinite interpretations are available.

(48)   [Context: A's officemate comes back to the office, after seeing the dean to ask for the permission to organise a conference, which A knows about.]
       A: __   *zgodziła*   *się?*
           (she)   agreed.SG.F   SE
           'Has she agreed?'

(49)   [Context: A and B see a roe deer approaching their house.]
       A: *Pod*   *dom*   *Anny*   *też*   *dziś*   __   *przyszła.*
           under   house   Anna.GEN   also   today   (she)   came.SG.F
           'It/one came near Anna's house today too.'

While these facts could be surprising when viewed through the lens of null subjects painted in the literature on Romance and Germanic languages, they are less surprising when they are viewed against the background of the broad interpretive possibilities of bare lexical NPs and overt pronouns in Polish, both of which can also be associated with the definite and the indefinite readings (see Ruda 2021a). This is illustrated for lexical NPs in (50), where the NP can receive either the definite or the indefinite interpretation, and in (51)–(52) for pronouns, where in the former case the pronoun is definite but in the latter it has the unspecific indefinite reading.

(50) *Pies   zaszczekał.*
　　 dog　barked.sg.m
　　 'The/a dog barked.'

(51) *Paweł spotkał　dziś　interesującą dziewczynę, ale* __ *nie*
　　 Paweł met.sg.m today interesting　girl　　　 but (he) not
　　 zaprosił　jej　na kawę.
　　 invited.sg.m her.gen on coffee
　　 'Paweł met an interesting girl today, but he didn't invite her for coffee.'

(52) A: __ *nie mam　dziewczyny.*
　　　　(I) not have.1sg girl
　　　　'I do not have a girlfriend.'
　　 B: *Czemu* __ *sobie　jej　nie znajdziesz?*
　　　　why　(you) self.dat her.gen not find.2sg
　　　　'Why won't you find one for yourself?'

If lexical NPs and overt pronouns enjoy such interpretive freedom in Polish, there is no reason why null pronouns should behave differently, barring some extant theoretical assumptions (in particular, [D] in T; see Holmberg 2005, 2010 and Roberts 2010, 2019). The hypothesis that the vP/VoiceP-internal subject is represented as a minimal pronominal element, that is n(P), which differs from lexical NPs (qua nPs) only in the lack of a lexical root (see Barbosa 2019, Ruda 2018, and for a more detailed theoretical proposal in relation to Polish, Ruda 2020) thus seems worth developing in future research. The final interpretation of such a subject *n*P can be taken to be achieved with the same mechanisms which yield the relevant interpretations in (50)–(52).[11]

　　　To sum up, the discussion in this section has shown that null subjects in Polish need not be interpreted with reference to an Aboutness-Shift Topic and their interpretation is not constrained by the presence of this topic in the overtly realised structure. Moreover, null subjects in Polish can be interpreted not only as definite, but also as indefinite, though they do not have quantificational readings. This is similar to what we observe with bare lexical NPs and with overt pronouns, warranting the postulated nP representation.

---

11 The ultimate interpretation of bare NPs in languages such as Polish can be derived by appeal to different semantic and pragmatic mechanisms, including, but not restricted to, (ɩ) type-shifting (for relevant discussions, see Chierchia 1998; Heim 2011; Šimík & Demian 2020, submitted; Šimík forthcoming and references therein).

5 Conclusion and further theoretical remarks

According to some recent approaches to third person null subjects in consistent NSLs, the licensing and interpretation of these subjects relies on the Aboutness-Shift Topic, represented in the clause overtly or covertly (see, in particular, Frascarelli 2007, 2018; Frascarelli & Jiménez-Fernández 2019; Holmberg et al. 2009; Holmberg 2010; Sigurðsson 2011, 2014, 2019 and Jiménez-Fernández 2016). The first goal of this contribution has thus been to show that this hypothesis does not generalize to all NSLs, its predictions not being fulfilled in Polish.

Apart from these empirical issues, there is a number of general problems with the Topic Criterion, if viewed through the lens of explanatory adequacy (see also Cole 2010 and Sigurðsson 2019 for some critique). More specifically, the condition requiring the linking of the null subject with a particular type of topic does not seem to follow from any deeper properties of natural language grammar or its interfaces. Also, as agreement has no role to play under this approach, it leaves unexplained why languages such as Italian do not make object drop available to the same extent and under the same conditions as null subjects. Furthermore, this approach stipulates that Aboutness-Shift Topics (both when continued and when newly introduced) and G(iven)-Topics can be silent. This silence is not explained in any way in relation to any current approaches to (deep or surface anaphoric) ellipsis. Since this mechanism is not explained, the nullness question is just shifted from the pronominal subject to the topics, or, more precisely, it is added on top of it, as the null status of *pro* linked with the topics needs to be accounted for, too. While with respect to G-Topics Frascarelli (2018) suggests some relation to copy deletion, this is highly unlikely: there need not be c-command between the assumed copies and no movement (or Roberts-style incorporation) is involved. Similar issues pertain to the formal nature of the assumed relation linking the Aboutness-Shift Topic, G-Topic(s) and *pro*. It seems that the intended relation is Agree, but here we face both the locality issue, as the linking needs to be able to operate across sentential boundaries, and the c-command issue, when the Aboutness-Shift Topic is introduced inside a phrase equivalent to *as for XP*. Overall, what this implies is that, even if this approach were not challenged by the empirical data discussed here, it would still be missing a complete account of the null realisation of the subject and the interpretation questions posed by it in consistent NSLs (and beyond).

Finally, on a more positive note, I would like to point out again that the discussion in this contribution also indicates that the investigations of the properties of null subjects in a language can benefit from considering them against the background of the interpretive properties of other nominal elements in the language, where the availability of definite and indefinite interpretations with bare

lexical NPs and pronouns may prove to be crucial to our understanding of the linguistic variation in the domain of argument drop.

References

Barbosa, Pilar. 2019. *Pro* as a minimal *n*P: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50. 487–526.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6. 339–405.

Cole, Melvyn Douglas. 2009. Null subjects: A reanalysis of the data. *Linguistics* 47(3). 559–587.

Cole, Melvyn Douglas. 2010. Thematic null subjects and accessibility. *Studia Linguistica* 64(3). 271–320.

Dalmi, Gréte. 2017. Little *pro*'s, but how many of them? – On 3sg null pronominals in Hungarian. *LingBaW* 3. 61–73.

Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*, 211–239. Oxford: Oxford University Press.

Frascarelli, Mara & Ángel L. Jiménez-Fernández. 2019. Understanding partiality in *pro*-drop languages: An information-structure approach. *Syntax* 22(2–3). 162–198.

Haider, Hubert. To appear. A null theory of scrambling. In Volker Struckmeier & Andreas Pankau (eds.), Special issue of *Zeitschrift für Sprachwissenschaft*. Berlin: Mouton De Gruyter.

Heim, Irene. 2011. Definiteness and indefiniteness. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, 996–1025. Berlin: Mouton de Gruyter.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97.

Holmberg, Anders & On-Usa Phimsawat. 2017. Truly minimal pronouns. *Revista Diadorim* 19. 11–36.

Jiménez-Fernández, Ángel Luis. 2016. When discourse met null subjects. *Borealis: An International Journal of Hispanic Linguistics* 5. 173–189.

Kibort, Anna. 2001. The Polish passive and impersonal in Lexical Mapping Theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG01 Conference*, 263–283. Stanford: CSLI Publications.

Kibort, Anna. 2008. Impersonals in Polish: An LFG perspective. *Transactions of the Philological Society* 106. 246–289.

Krzek, Małgorzata. 2010. Some aspects of subjects of impersonal constructions in Polish. *Newcastle Working Papers in Linguistics* 16. 66–87.

Krzek, Małgorzata. 2013. Interpretation and voice in Polish SIĘ and -NO/-TO constructions. In Irina Kor Chahine (ed.), *Current Studies in Slavic Linguistics*, 185–198. Amsterdam: John Benjamins.

Krzek, Małgorzata. 2017. The null subject parameter meets the Polish impersonal -NO/-TO construction. In Michelle Sheehan & Laura R. Bailey (eds.), *Order and Structure in Syntax II: Subjecthood and Argument Structure*, 307–318. Berlin: Language Science Press.

Lavine, James L. 2005. The morphosyntax of Polish and Ukrainian *-no/-to*. *Journal of Slavic Linguistics* 13. 75–117.

Lavine, James L. 2017. Syntactic change and the rise of transitivity: The case of the Polish and Ukrainian *-no/-to* construction. *Studies in Polish Linguistics* 12. 173–198.

McShane, Marjorie J. 1999. The ellipsis of accusative direct objects in Russian, Polish and Czech. *Journal of Slavic Linguistics* 7. 45–88.

McShane, Marjorie J. 2009. Subject ellipsis in Russian and Polish. *Studia Linguistica* 63. 98–132.

Nillson, Barbro. 1982. *Personal Pronouns in Russian and Polish: A Study of Their Communicative Function and Placement in the Sentence*. Stockholm: Almqvist & Wiksell International.

Pisarkowa, Krystyna. 1969. *Funkcje składniowe polskich zaimków odmiennych* [Syntactic functions of the Polish inflecting pronouns]. Wrocław: Zakład Narowowy Imienia Ossolińskich/Wydawnictwo Polskiej Akademii Nauk.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar. Handbook in Generative Syntax*, 281–337. Dordrecht: Kluwer.

Roberts, Ian. 2010. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 58–87. Cambridge: Cambridge University Press.

Roberts, Ian. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Ruda, Marta. 2014. The impersonal subject *-N/T* construction in Polish and the typology of Voice heads. *Studies in Polish Linguistics* 9. 203–243.

Ruda, Marta. 2018. Syntactic representation of null arguments: [$_{NP}$ e] as [$_{nP}$ n] (minimally). In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*, 240–260. Oxford: Oxford University Press.

Ruda, Marta. 2020. Analyses of null arguments meet a bare NP, consistent NS, partial NO language, handout, *9th Cambridge Comparative Syntax Conference*, Cambridge, UK (online), September 2020.

Ruda, Marta. 2021b. Strict and sloppy readings of pronominal objects in Polish. *Studies in Polish Linguistics* 16. 121–144.

Ruda, Marta. 2021a. Polish personal pronouns: [$_{PersP}$ Pers [$_{NumP}$ Num [n]]] and [$_{NumP}$ Num [n]]. *Studies in Polish Linguistics* 16. 23–40.

Samek-Lodovici, Vieri. 1996. *Constraints on subjects: An optimality theoretic analysis*. PhD dissertation. Rutgers University, New Brunswick, NJ.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42. 267–304.

Sigurðsson, Halldór Ármann. 2014. Context-linked grammar. *Language Sciences* 46. 175–188.

Sigurðsson, Halldór Ármann. 2019. Topicality in Icelandic: Null arguments and narrative inversion. In Valéria Molnár, Verner Egerland & Susanne Winkler (eds.), *Architecture of Topic*, 249–272. Berlin: Mouton De Gruyter.

Šimík, Radek. Forthcoming. Inherent vs. accidental uniqueness in bare and demonstrative nominals. In Uwe Junghanns, Andreas Blümel, Jovana Gajić & Hagen Pitsch (eds.), *Advances in Formal Slavic Linguistics* 2018. Berlin: Language Science Press.

Šimík, Radek & Christoph Demian. 2020. Definiteness, uniqueness, and maximality in languages with and without articles. *Journal of Semantics* 37. 311–366.

Šimík, Radek & Christoph Demian. Submitted. Uniqueness and maximality in German and Polish: A production experiment. In Mojmír Dočekal & Marcin Wągiel (eds.), *Formal Approaches to Number in Slavic and Beyond*. Berlin: Language Science Press.

Willim, Ewa. 1989. *On Word Order: A Government-Binding Study of English and Polish*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Ludmila Veselovská

5 The features of null subjects: A case study of Czech

1 Czech, a consistent Null Subject Language

Since classical linguistic studies it has been noticed that some parts of the assumed clausal structure can be omitted (dropped) in contexts where they are semantically or pragmatically inferable. The variety of recoverable omissions includes the subject of the clause. This language-specific characteristic has been widely discussed in classical and modern linguistics because both Ancient Greek and Latin represent examples of such languages, as opposed to modern, mostly Western European languages. With the development of current linguistic frameworks, *pro*-drop or null subject (NS) phenomena remain an important part of comparative studies and are treated with the use of a range of framework-specific terminologies and methodologies.

Perlmutter (1971) brought the concept of *pro*-drop/NS first into mainstream generative grammar within the Extended Standard Theory. He discussed some general characteristics of Null Subject Languages (NSLs), comparing them with those of English. For example, his initial study introduced the correlation with the *that*-trace effect, which is active in non-*pro*-drop English but not attested in the NSLs – e.g. Czech – as illustrated below in (5).

In the Government & Binding (GB) framework, Chomsky (1981, 1982) analysed NS structures to introduce a new syntactic entity – the covert (invisible, phonetically empty) subject of finite predicates (little *pro*) into his taxonomy of nominal expressions.[1] To explain its distribution, Chomsky used the binary [±P: pronominal] and [±A: anaphor] features, which represent the structural requirements imposed on various nominals.

The four types of nominal expressions in (1) show the possible combinations of the ± values for the A/P features as suggested in Chomsky (1981). In this taxonomy, each type of nominal element was represented with a phonetically realized (overt) variant and a phonetically unrealized (empty, covert) one, as well.

---

1 The null subject of finite clauses is labelled "little *pro*" to be distinguished from the "big PRO", which represents the empty subject of infinitives, as in the [+A, +P] line in (1).

(1) Four types of pronominal expressions in the GB framework

Features	Overt	Covert
[-A, -P]	a nominal (R-) expression	Wh-trace (t_{wh})
[-A, +P]	a personal pronoun	little *pro*
[+A, -P]	an anaphoric pronoun	NP-trace (t_{NP})
[+A, +P]	feature conflict: no overt realization	big PRO

GB studies collected relevant data for a wider variety of languages (especially Romance and Germanic), defining the representative cross-linguistic concept of the Null Subject Parameter (or *pro*-drop Parameter). This parameter became one of the most general morpho-syntactic characteristics used to compare language systems.

In mainstream GB studies,[2] NSLs were defined as languages which set the *pro*-drop Parameter in such a way that they have the null pronoun *pro* in the subject position. However, since the beginning it was not only the lack of overt subjects (i.e. the presence of *pro*) that represented the parametric option. Apart from testing the extent which allows a language to use some variety of *pro* (in some contexts), the studies also concentrated on looking for other phenomena which correlate with the availability of an empty pronoun *pro*. By the end of the 1980s, the following list of characteristics was assumed to form a parametric cluster:

(2)   *Pro*-drop/NS-parametric cluster[3]
   a. existence of *pro* – a silent null subject of a finite clause;
   b. no active *that*-trace effect;
   c. both pre- and post-verbal overt subjects possible (free subject inversion);
   d. rich (alternatively, "uniform") inflection across person and number specifications;
   e. a definite (or specific) interpretation of 3SG *pro*/NS (marked *pro*/NS_arb).

Some Slavic languages have always been classified as *pro*-drop languages, and Czech may serve a relevant picture at this point (see the descriptive summary

---

2 See Rizzi (1982, 1986) and Jaeggli & Safir (1989). For a more framework-free discussion, see Haspelmath (2001) and Haspelmath & Dryer (2005), among many others.
3 The relevance of the availability of overt post-verbal subjects was discussed for Romance languages in Kayne (1980, 1989). The interpretation (its non-specificity) is described as typical for some but not all NSLs in Gilligan (1987).

in Short 1990). The examples in the next section demonstrate that all the above characteristics of a NSL can be illustrated in Czech.

1.1 Parametric cluster characteristics in Czech

Regarding the attested NS (the existence of a *pro*) mentioned in (2a) in the preceding section, some languages allow the use of *pro* only in some contexts: e.g. in some tenses or some persons, in (some) subordinated clauses, or only with special interpretation. As for Czech, there are only minimal restrictions on the use of NS.

   The examples in (3) illustrate that *pro* can be used in any person/number combination in (3a) and in any tense/aspect combination in (3b).[4] We can also see that any of these clauses can be either root or embedded.

(3) *Petr   si   myslí,   že...*
     Peter.NOM RFL   think.3SG   that
     '(Peter thinks that)...'
     a. *...spím   /spíš   /spí   /spíme   /spíte.*
         sleep.1SG sleep.2SG sleep.3SG/PL sleep.1PL sleep.2PL
     b. *...jsem   (za)spala   (za)spím   budu*
         be.1SG   (PFX)sleep.PTCP.SG.F   (PFX)sleep.PRS.1SG FUT.1SG
         spát.
         sleep.INF
         '...I (over)slept / (over)sleep / will sleep.'

The covert (null) subjects in (3) are interpreted as referential arguments of the predicate and they can also be expressed by overt pronouns, when pragmatically prominent (see the discussion in Section 1.2). The subject in (4) is an expletive, i.e. non-referential. It triggers a 3rd person singular neuter (default) agreement and it usually does not have any overt realization in Czech. Example (4) features

---

4 The glosses in (3b) contain the labels PST/PRS/FUT (past/present/future). However, I do not claim any specific tense morpheme or tense features to be present in these Czech examples. Instead, I assume a more analytic (combinatorial) strategy leading to the traditionally labelled interpretation of PST/PRS/FUT, which includes mood. Further on, in this chapter I will mark the *l*-participle (the verbal form which appears as a part of the Czech analytic past tense and present/ past conditional) as PTCP plus number and gender features (i.e. not providing the tense feature) in the glosses. The English translation provides the interpretation and, as for the discussion of the precise feature content of the *l*-participle (not relevant for the discussion in this study), I refer interested readers to Emonds & Veselovská (2015).

predicates in past, present and future(s), which serves to show that *pro* is available in Czech in all tenses. [5]

(4) (*Myslím, že*)   *(na)pršelo*   *(na)prší*   *bude*   *pršet.*
    think.1SG that   (PFX)rain.PTCP.3.SG.N (PFX) rain.3SG FUT.1SG   rain.INF
    '(I think that) it rained / is raining / will rain.'

In the above parametric cluster, in (2b), Perlmutter's *that*-trace effect is mentioned as another typical characteristic found in NSLs. In (5), we can see the contrast between the Czech example and the English translation. In English, when a subject WH pronoun *who* is extracted from the embedded clause to the sentence initial position, the complementizer *that* must not be present. The Czech equivalent is fully grammatical – the complementizer (most likely *že* 'that' subcategorized by the matrix verb) is in fact obligatory.

(5) *Kdo si myslíš   *(že) to   přinesl?*
    who RFL think.2SG that it.ACC bring.PTCP.(3)SG.M
    'Who do you think (*that) brought it?'

The *that*-trace filter distinguishes long distance extraction out of subject position from long distance extractions from V-complement and adjunct positions. This distinction is typical for the most overtly configurational languages that show hierarchical superiority effects. In Czech (as in many other Slavic languages) the effects of superiority are more difficult to find.[6]

The next characteristic (2c) in the parametric cluster is the position of overt subjects, and it was widely accepted after the discussion by Rizzi (1982), who provides data from Italian and proposes that the availability of *pro* in a given language correlates with the existence of post-verbal subjects, the so-called "overt subject inversion". In (6) we can see the Czech equivalents of his examples. (6a) and (6b) illustrate a DP subject and a pronominal one, respectively. Ignoring the distinctions in pragmatic interpretation, which are characteristic of the presence

---

5 As discussed in detail in the following sections, the features of person, number and gender never co-occur in the morphology of a single Czech verbal form: the features of person and number are visible on the predicate in present tense. The gender is marked only on the *l*-participle in past where the 3rd person auxiliary is zero. Therefore, I will mark the 3rd person on the *l*-participles in brackets only as (3) to signal that the 3rd person feature is *not* present there but it is interpreted only because of the zero (lack of the) auxiliary. In fact, also singular and masculine are only interpreted and have no overt morphological realization.
6 For superiority in Slavic, see Bošković (1997, 2011), Meyer (2004) or Shuurman (2016).

and position of the subject,[7] the examples nevertheless show that, generally speaking, a Czech subject can be dropped and when present, it can both precede and follow the verb.

(6) a. (*Náš*   *Petr*)   *dopis*   *nenapsal*   (*náš*
         our.NOM   Peter.NOM   letter.ACC   not.write.PTCP.(3)SG.M   our.NOM
         Petr)
         Peter.NOM
         'Our Peter did not write the letter.'

     b. (*já*)   *doma*   *pracovat*   *nebudu*   (*já*)
         I.NOM   at.home   work.INF   not.FUT.1SG   I.NOM
         'I won't be working at home.'

The parametric characteristic of NSLs in (2d) mentions "rich morphology". This property has been widely discussed already in classical, traditional and functionalist frameworks in order to account for the availability of *pro*. In terms of the concept of recoverability, the phonetically empty subjects are possible because their understood content can be inferred – in other words, because they are "identified" with the bound morphology present on the finite predicate.

   The notion of rich inflection has been adopted from traditional grammars into the generative framework in Taraldsen (1980) and has been under discussion ever since. However, a more precise definition of morphological richness, has never been that easy to provide. First, this is because of the failed attempts to list the (number of) specific features which are either sufficient or necessary to allow NS. This may signal that the traditional repertory of agreement features (if such features exist at all) is not suitable for the definition. Moreover, the proposals have to face the problems of zero bound morphemes, default and unmarked forms, and the fused nature of the inflectional morphology typical of Indo-European *pro*-drop languages. These factors resulted in the hypothesis of Jaeggli & Safir (1989), where the notion "rich" was replaced by a comparably vague concept of "uniform". Reference to the richness or morphological uniformity has become even more problematic to implement in the analysis of Asian *pro*-drop languages which do not have any agreement morphology at all.

   Leaving aside the problematic definitions of the required feature content, which may well be language-specific, Czech morphology is for sure both suffi-

---

7 In a NSL, an overt pronominal subject is always marked in some way. As I will mention several times in the study below, overt pronominal subjects in Czech, especially in the clause final position, are marked for contrastive focus.

ciently rich and uniform, although its verbal paradigms arguably also contain zero morphemes in predictable contexts (i.e., in present tense, 3rd person singular forms). Some aspects of the Czech verbal morphology are described in the sections to follow. In (7), three most frequent present tense verbal paradigms are presented in order to show that person and number of the subject are always recoverable from the inflection.[8]

(7) Czech verbal morphology in present tense

	'bring' (1st V-class)	'speak' (3rd V-class)	'do' (5th V-class)
1SG	*nesu*	*mluvím*	*dělám*
2SG	*neseš*	*mluvíš*	*děláš*
3SG	*nese*	*mluví*	*dělá*
1PL	*neseme*	*mluvíme*	*děláme*
2PL	*nesete*	*mluvíte*	*děláte*
3PL	*nesou*	*mluví*	*dělají*

As for gender, that can be seen only in the adjectival type of agreement on the two past and passive participles.

1.2 Interpretation of NSs

The last characteristic (2e) in the parametric cluster of NSLs above mentions a definite (specific) interpretation of 3SG *pro*/NSs and marked *pro*/NSs with arbitrary interpretation. In other words, NSLs are supposed to lack 3SG non-specific (generic) *pro* subjects.

To consider the Czech data with respect to this characteristic, let us first look at the traditional grammar analysis of Czech clauses without overt subjects.[9] The grammar textbooks distinguish (a) two-member clauses with "unexpressed" (covert) subjects, and (b) one-member clauses which entirely lack subjects. Such unexpressed subjects are illustrated in (3), (5) and (6) above. A so-called one-member clause has already been shown in (4), and some others are in (8) below. Notice

---

8 Traditionally, the complex system and taxonomy of Czech verbal paradigms include 5 classes with several sub-classes, the distinctions between which can be minimal and often phonetically motivated. None of the paradigms, however, disallows NS structures, and therefore I will ignore the variety here. For such lists and examples, see Grepl & Karlík (1998) or Short (1990).

9 For terminology and descriptive generalizations see the Czech grammar manual of Grepl & Karlík (1998). In this study I owe many of the Czech examples to personal communication with Petr Karlík.

that *weather*-verbs or "psychological" predicates in these one-member clauses have an unmarked (default) 3rd singular neuter agreement. The complements of these verbs take the form of adverbs, some contain a dative candidate for subject-hood and some structures are reflexive.

(8) a. *Je   chladno.*
       is.3SG   cold.ADV
       'It is cold.'

    b. *Bylo   mi   smutno.*
       be.PTCP.(3)SG.N   me.DAT   sad.ADV
       'I was sad.'

    c. *V kamnech   praskalo.*
       in stove   crackle.PTCP.(3)SG.N
       'It was crackling in the stove.'

    d. *Petrovi   se   dýchá   dobře.*
       Peter.DAT   RFL   breathe.3SG   well.ADV
       'It is easy for Peter to breathe.'

A non-specific and [+human] generic interpretation is typical for impersonal reflexive structures like (8d) or (9). The reflexive clitic *se* combines with the 3rd singular neuter agreement (the person and number features are visible in the present tense, and the number and gender in the past participle).[10]

(9) a. *Tancuje   se   každý den.*
       dance.3SG   RFL   every   day
       'One can dance every day.'

    b. *U večeře   se   mluvilo   jenom   francouzsky.*
       at dinner   RFL   speak.PTCP.(3)SG.N   only   French
       'Only French was spoken during the dinner.'

    c. *V Anglii   se   jezdí   vlevo.*
       in England   RFL   drive.3SG   left
       'In England they drive on the left.'

With agentless predicates (e.g., *weather*-verbs), NS is common is Standard Czech. The presence of an overt subject is rare but still attested. Such expletive subjects

---

10 Czech impersonal passive is described in detail in Hudousková (2009), who assumes that the subject of impersonal passives is represented by the reflexive pronoun (a deficient nominal head) with arbitrary [+human] interpretation. Generic interpretation with respect to the derived imperfective morphology is discussed in Filip & Carson (1997).

take the form of the 3rd person singular pronoun *ono* (glossed as 'it' in (10)) or a weak form of a demonstrative *to* (glossed as 'that' in (10)). Two examples used above are modified below to show that the equivalent of the NS is more likely to be the pronoun *ono* 'it'. As for the demonstrative, its presence modifies the meaning and sometimes it can even combine with the 3rd person neuter personal pronoun *ono* 'it'.

(10) a. *To   je   chladno.*
         that   is.3SG   cold
         'It is cold, isn't it?'
     b. (*Ono to / To to*)   *v*   *kamnech*   *praskalo.*
         it that / that that   in   stove   crackle.PTCP.(3)SG
         'It was crackling in the stove.'
     c. *Ono/To   se   tancuje   každý den.*
         it/that   RFL   dance.3SG   every day
         'One can dance every day.'

Assuming that Czech has to have a discourse-interpreted Topic in the CP domain, the demonstrative pronoun (in nominative) is more likely to represent a Topic discourse particle.[11] Their feature compatibility with the default verbal agreement is, however, robust.

The NS parametric cluster in (2e) above mentions a definite (specific) interpretation for the 3rd singular *pro*/NS. Other person specifications were not mentioned and, indeed, in Czech a non-specific generic interpretation can be obtained with clearly two-member sentences containing NS – if the context requires and/or allows it – with *pro*/NS of other person and number specifications. Some of the structures have idiomatic interpretation.

The examples below show that both 1st and 2nd person agreement is acceptable for a non-specific interpretation. The generic interpretation in (11a–c) is inclusive (i.e. it includes both the speaker and the hearer).[12]

---

11 In (10b) we can see that the weak demonstrative expletive *to* ('that') can combine with the indefinite pronoun *ono* ('it') or it can be doubled. Both combinations add a feeling of surprise or emphasis related to the whole proposition. Fanselow & Lenertová (2001) and Šimík (2007), among others, refer to the concept of the Focus head and the left-hand Focus to analyse the demonstrative *to*...

12 The examples in (11a-c) are well known traditional Czech proverbs (idioms). In the non-idiomatic examples, 2nd singular would probably be the most common form. On the other hand, I did not find a 2nd plural with an inclusive reading.

(11) a. *Když   něco   slíbím   tak   to   i   splním.*
      when   something.ACC promise.1SG then   it.ACC also   fulfil.1SG
      'When I promise something, I also fulfil it.'

      b. *S   poctivostí nejdál   dojdeš.*
      with   honesty   furthest   go.2SG
      'You get furthest when you are honest.'

      c. *Lilek   zapékáme   20 minut.*
      eggplant   roast.1PL   20 minutes
      'We roast the eggplant for 20 minutes.'

As for the 3rd person, the singular of which is mentioned explicitly in (2e) of the NS parametric cluster, example (12) shows the Czech agreement for the 3rd plural *pro*. In (12) the speaker and hearer are excluded, but the generic non-specific interpretation is allowed (in a suitable context and restricted/implied with the locational adverbial).

(12) *V   Česku   neprodávají / neprodávali   skopové.*
      in   Czechia   not.sell.3PL / not.sell.PTCP.(3)PL.M   mutton
      'They do/did not sell mutton in Czechia.'

Example (13a) is the repetition of the non-idiomatic (11c) with the agreement for the 3rd person singular *pro*. In turn, (13b) serves to demonstrate that, in accord with (2e), the non-specific interpretation is *not* allowed for the Czech 3rd singular *pro*.

(13) a. *Lilek   zapéká   / zapékal   20 minut.*
      eggplant.ACC   roast.3SG / roast.PTCP.(3)SG.M   20 minutes
      'He/*One roasts/roasted eggplant for 20 minutes.'

      b. *Každý den potkává / potkával   nové kamarády.*
      every   day   meet.3SG / meet.PTCP.(3)SG.M   new   friends.ACC
      'He/*One meets/met new friends every day.'

If the nonspecific subject is to trigger the 3rd person singular agreement, it has to be with the overt, idiomatic and a somewhat obsolete pro-form *jeden* 'one'. As it is the case with its English translation equivalent *one* used in the gloss, the pragmatic context decides whether (14) is interpreted as an indefinite pronoun or a numeral (in Czech numerals do not have to be followed by a nominal complement).

(14) a. *Jeden   nikdy neví.*
      one.NOM   never   know.3SG
      'One never knows.' / 'One of them never knows.'

    b.   *Jeden*   *se*   *nestačí*   *divit.*
          one.NOM   RFL   not.mange.3SG   marvel.INF
          'One (of them) can't help being surprised.'

In all the examples above the glosses suggest that the overt pronominal subjects have the same φ-features as their NS counterparts, as shown by the agreement morphemes in table (7). Therefore, using overt pronouns together with the agreement morphemes may seem unmotivated, representing a kind of undesirable or uneconomical φ-feature doubling.

    In Emonds (1987), a specific version of the GB framework, the φ-features of predicate agreement represent an Alternative Realization (i.e. an equivalent) of the pronominal subject. Emonds's Invisible Category Principle then predicts that economy excludes one of the variants if they are fully equivalent. If this holds, some formal and definable distinction between overt subjects and NS is therefore theoretically desirable, and – indeed – a comparison of the interpretation(s) of overt and covert subjects is never ignored in the discussions of NS.

    Given that the distinction between overt and null subjects cannot be easily expressed in terms of φ-feature content, it has to be expressed using some other kind of features. Most analyses of NS discuss the distinction in terms of discourse concepts, referring to specificity, the theme/rheme distinction, Topic/Focus realization, etc.

    To illustrate the discussion using the Czech data from this section, examples (11) and (12) demonstrate NS with non-specific readings. If the subjects were realized as overt pronouns, the interpretations would have to refer to some specific participants in a specific discourse. Moreover, with the 1st and 2nd person pronouns, which are interpreted with respect to the speech act, the overt forms contain (apart from the φ-features) a discourse feature of Focus or contrastive stress. In Czech these pronouns are typically found in clause peripheral positions (both initial and final) and/or are phonetically prominent.

    As for the 3rd person singular form NS in (13), it carries only a specific anaphoric interpretation. Similar specific interpretation is also required with anaphoric NS in clauses with a hierarchically symmetric interpretation. In these contexts, the overt pronoun would represent a change of reference of the Topic (subject). (In English, the equivalent of the overt pronoun is represented by a stressed pronominal.)

(15)   *Petr*$_i$ *pozdravil*   *Karla*$_j$ *a*   pro$_i$ / *on*$_{*i/j/k}$
     Peter.NOM greet.PTCP.(3)SG.M Charles.ACC and (3SG) / he
     odešel.
     leave.PTCP.(3)SG.M
     'Peter greeted Charles and he/HE left.'

Examples like (15) suggest that while overt pronominals are a kind of fully specified structure and therefore truly syntactically free in accord with Principle B (Chomsky 1981: Chapter 3), the NS/agreement cluster of φ-features may require some additional relation to a specific discourse feature, plausibly realized in the CP domain of the same clause.

Abstracting away for a moment from the mainly discourse-based distinction between overt and null subjects, it appears that most Czech structures allow either of them. To my knowledge, there is only one exceptional type of clause that does not allow a NS in Czech regardless of the interpretation, i.e. the "pointing structures" with copulas (those with an initial demonstrative and a nominal predicate following the verb *be*). In example (16a) below, we cannot omit the initial demonstrative *to* (the 3rd person singular), plausibly representing the Topic pointed to and formally marked most likely by nominative case (the form is syncretic with accusative). In turn, (16b) shows that neither can the DP following the copula be omitted. This DP is represented by a pronoun *já* 'I' in (16) to show that it also carries nominative and, moreover, agrees with the copula.[13]

(16)   a.   *To*/*Ono*/*pro   *jsem*   *já.*
           that/it/(it)   be.1SG   I.NOM
           'It is me.'
       b.   *To*   *jsem*   *já*/*pro.
           that   be.1SG   I.NOM/(I.NOM)
           'It is me.'

The explanation of the unacceptability of NS in the pointing structures may be pragmatic – the agreement present on the copula does not reflect the initial demonstrative and, therefore, the NS replacing this demonstrative in (16a) cannot be licensed by the agreement and is ungrammatical. On the other hand, the nominal predicate represents new information and the NS can never be interpreted as a stressed (rhematic, focused) constituent.

---

13 In the Czech structure DP copula DP, both the DPs are nominative. As for agreement, the copula picks the φ-features of one of the two DPs – those which are more marked. In the pointing structures as in (16), the demonstrative is the unmarked 3rd singular neuter form and, therefore, the features reflected by the copula are the features of the second DP – the marked 1st person singular. The system of agreements with the copula in Czech is discussed in detail in Bartošová (2017).

2 Minimalism and micro-parameters

In the preceding section, we saw that Modern Czech shows all the characteristics required to be labelled as a full *pro*-drop or consistent NSL, in accordance with the criteria set out in (2).

Using the standard GB framework, Rizzi (1986) analysed the *Pro*-drop/ NS Parameter as a complex of two parametric characteristics. He attempted to explain all the attested correlations summarised in (2) referring to the then recent Empty Category Principle. According to the ECP, all empty categories, including the NS *pro*, need to be both structurally licensed and identified. The two requirements are stated in (17).

(17)   ***Pro-drop Parameter*** (Rizzi 1986: 518–523)
    a.   Licensing: *pro* is case-marked by X^0_y – where y is parametrized.
    b.   Identification: *pro* inherits the f-feature values of X^0_y, if it has f-features; if not, *pro* gets a default interpretation (typically arbitrary).

The f-features are [±Pron] and [±Ref]. In (17a) Rizzi proposes that structural government is needed to license *pro* in a context (i.e. a case-assigning head X^0), and in (17b) he states the identification conditions, which include some setting for non-specific f-feature values.

In the 1990s, the *pro*-drop/NS Parameter in the form of (17) underwent some rather substantial changes in order to become compatible with the development of the generative framework in its minimalist stage. Even more importantly, the theory had to face some more controversial empirical data brought into light by Gilligan (1987), among others. Providing a more detailed and focused description of a typologically wide variety of languages, including non-European ones, Gilligan (1987) checked the validity of the list of the characteristics in (2) and showed that in spite of the existence of some correlations between the properties, not all NSLs share all of the characteristics as one uniform cluster.[14] Based on his data analyses, Gilligan proposed a distinction between partial NSLs and full NSLs. He also argued that the two groups systematically differ in the form of default agreement with thematic and non-thematic NS. Specifically, partial NSLs allow a non-thematic NS, while full NSLs allow only thematic NS.

---

14 The data showing variation in the NS characteristics for Slavic languages (Russian, Polish and Czech) had earlier been provided and analysed in Franks (1995).

2.1 Features and functional heads

In the early version of the Minimalist Program, Chomsky (1991, 1993) stated that most syntactic operations (including agreement and movement) were related directly to the feature characteristics of the non-lexical, i.e. functional heads. The feature content of these heads (their strong vs. weak characteristics) was arguably subject to diachronic and cross-linguistic variation. This theoretical assumption motivated further reformulations of the already dual NS Parameter in (17) and the introduction of several interconnected but independent micro-parameters.

The current micro-parametric approach to NSLs is rooted in the studies by Alexiadou & Anagnostopoulou (1998), Holmberg (2005), Alexiadou (2006), and it is represented by the studies in Biberauer, Holmberg, Roberts & Sheehan (2010) and in other studies cited in D'Alessandro (2015). The authors accept the so-called Borer-Chomsky conjecture, which states that cross-linguistic variation should be limited to the parametric choice of the features located on plausibly universal functional head(s) (see Borer 1984). For NS/*pro,* the functional heads are those related to the syntactic position of clausal subjects.

The functional head relevant for NS (located in the high periphery of the extended verbal projection) will be labelled here as T (the labels used in other studies include INFL/AGR^S or I). In the proposed analysis, the head T – apart from its own interpretable categorial feature T (related to case) – can (in some languages) host two other feature sets, namely [φ] and [D].[15]

The parameter referring to *pro*-drop/NS characteristics is then explained as a result of the combination of interpretable ([φ]/[D]) and uninterpretable ([*u*φ]/[*u*D]) features of the functional head T (the Probe) and a nominal element in the position of subject (the Goal).

Roberts & Holmberg (2010)[16] propose that the relevant Probes can host (fully or partially) *specific* vs. *impoverished* values of the φ-features. The unvalued (strong, uninterpretable) [*u*D]/[*u*φ] of the Probe must be valued in relation to an available Goal. The Probe thus looks for the required [D]/[φ] features of some overt or covert nominal element in its c-command domain. Given that the [*u*D]/[*u*φ] features are located in T, the available relevant nominal element is the subject.[17]

---

15 As mentioned above, in Rizzi (1986) the earlier version of these features was marked as [±Pron] and [±Ref].

16 See also Roberts (2010) and Barbosa (2019).

17 Chomsky (2013) proposes that the "strength" of T in a full NSL is due to the language-specific property of its φ-features: the φ-features of T are strong if they are able to label the TP projection as φP (DP/TP=<φ,φ>).

In other words, the theory claims that the presence/absence and the strong/weak combination of those two features are the main source of the attested distinctions between the NSLs and non-NSLs and also between partial and consistent (full) NSLs. The proposed classification of NSLs is listed in (18).

(18) a. Radical/discourse NSLs
      a. Partial NSLs
      b. Consistent NSLs
      d. Non-NSLs, which do not allow NS

The radical/discourse NSLs are those that allow NS, but they have little or no agreement morphology at all. The detailed characteristics of these discourse (Topic) *pro*-drop languages are found in e.g. Huang (1984), Tomioka (2003), and Saito (2007). They will not be discussed here. The partial NSLs allow indefinite NS and their NS can appear only in a defined set of contexts, e.g., with some person or tense specifications or in embedded clauses. A typical example of a partial NSL is Finnish (see Holmberg 2005). As for Slavic languages, Russian is a good candidate. In turn, Czech is a perfect representative of consistent NSLs. Abstracting away from any potentially problematic parts of the micro-parametric approach to NS sketched above, I will use it to analyse the Czech data.

The feature set relevant for the consistent NSLs proposed by Roberts & Holmberg (2010) is schematically listed in (19). Notice that it refers to the Probe (presumably the functional head T) as well as to the Goal (the characteristics of a nominal constituent in the subject position). The factors that make this feature set language-specific are (i) the presence and (ii) the values of the features:

(19) Features of T/D in a consistent NSL
      a. Probe: T [$u\varphi$], [uD]
      b. Goal: [$_{DP}$ D$_{[\varphi]}$. . . [$_{n/NP}$ *one*]] or [$_{DP}$ D$_{[\varphi]}$. . . [$_{n/NP}$ pro]]

Assuming that Czech is a consistent/full NSL, (19a) states that the Czech functional head T contains both [φ] and [D] features and that both of them are in the strong/uninterpretable form, i.e. they need setting/valuation. As for the Goal, to comply with (19b), Czech nominal constituents must be DPs, i.e. they must include the D level. Veselovská (2014, 2018) provides multiple arguments for the claim that the DP analysis is compatible with and, in fact, strongly supported by Czech paradigms, in spite of the fact that Czech lacks overt determiners. Moreover, within the framework of Roberts & Holmberg, the nominal complexes

in a consistent/full NSL must have DPs with (i) a *rich* D and (ii) the n/NP comple-
ment serving as a *variable* bound by $D_{[\varphi]}$.[18]

2.2 A structure for agreement

The standard agreement relation used in the GB period was based on the Spec-
head relation, i.e. the relevant structure had a phrasal constituent in the Spec
position of the phrase projected by the agreeing head. With finite clause subjects,
the position was in the domain of the functional head T. The schematic picture
of a specifier-head agreement is presented in (20a) below. As for the structures
with no overt DP subject, Alexiadou & Anagnostopoulou (1998) argue that the
φ-features on the finite predicate do not represent an agreement of the head T
with the (overt or covert) subject in Spec,TP. They propose instead that NSLs exert
a kind of head-to-head (feature) movement, illustrated schematically in (20b).
The agreement morphology is a suffix, so the φ-features presumably incorporate
to the right of the T head.[19]

(20) a. Spec-head agreement   b. Feature movement
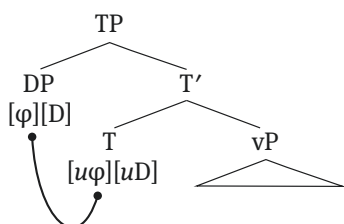
---
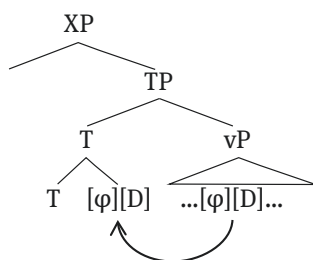
18 The characteristics of the φ-features with respect to the NS phenomena are also discussed in
e.g. Williams (1981, 1994); Longobardi (2008); Richards (2014); and Barbosa (2019). The authors
describe the anaphoric properties of the NS, which requires binding by speech-act features in C
or in a superordinate structure.

19 From the discussion in Alexiadou & Anagnostopoulou (1998), it is not clear whether the pro-
cess suggested for the agreement morphology licensing NS (cf. (20b)) also takes place (in a NSL)
with overt subject structures. As discussed later in the sections to follow, the morphology with
overt and covert subjects is identical, at least in Czech. Theoretically, however, the source of the
morphology should be distinct; while (20b) is the structure available in NSLs, (20a) alone would
not license the NS.

Roberts & Holmberg (2010) base their analysis of NS on the above proposal. They describe the checking of the [$u\varphi$] and [uD] features of T as a process which includes the following three steps.

First, the unvalued [$u\varphi$] of T is checked by a kind of movement of the subject's φ-features to T in a structure, as in (20b), i.e. the Goal (subject) features move to T and the Probe and Goal get unified under one label. However, Holmberg (2005) describes the process in (20b) not as a standard head or phrase movement, but as a sort of incorporation of a subject "clitic pronoun" into T, called "D-to-T incorporation". The incorporated subject clitic, then, is what is traditionally called "rich" inflection, illustrated here for Czech in the verbal paradigm in (7). After the incorporation, the lower part of the φ-feature "chain" (the subject's theta-position) can be deleted under feature identity.[20]

The second step in the valuation of the Probe features [$u\varphi$] and [uD] of T, stated above in (19), concerns the licensing of the subject reference with respect to the specific speech act conditions. Building on the proposal of Roberts & Holmberg (2010), I assume that the context features entering the relation are structurally represented – they are hosted in a functional head Topic (the "Aboutness Topic") located in the high clausal periphery.[21] I will show below that the discourse features have a role in the licensing (identification) of the person (and number) features.

The final step described in (19b) concerns the binding of the variable inside the n/NP. This variable has to get reference to be discourse bound by $D_{[\varphi]}$.[22] I propose in Section 4 that the binding process leads to the delayed unification of the φ-features of N- and D-domains. However, the unification need not be morphologically visible, as in the examples of certain φ-feature mismatches in Czech analytic predicates illustrated in Section 4.2.

---

20 As for this kind of movement, the higher part of the chain is a minimal (head) constituent (the inflectional morpheme). This may raise a question about the constituent characteristics of the other member of the assumed chain – the theta-position of the subject. Does the inflection represent the whole subject or only a part of it? Is the copy represented by the inflection identical with the copy of a covert element in the subject's base position? I assume that the term "clitic incorporation" is used to label the otherwise non-standard process – with no precise analysis of the phenomena. These questions go beyond the scope of this chapter and will not be resolved here.
21 See also Kučerová (2014) and Veselovská (2020).
22 For more discussion see Vikner (1995), Zanuttini & Bernstein (2008) and Kučerová (2007, 2012), as well as the contributions mentioned in footnote 18.

3 Agreement morphology in Czech

In inflectional languages, the valuation of uninterpretable strong features – including presumably clitic incorporation – is expected to be signalled with some overt morphology on the Probe itself or in a very local domain containing the Probe.[23] It is to be still decided to what extent and under what conditions the overt inflection correctly encodes (all of) the syntactic features and their values. In this study, I will assume the less controversial one-way correlation: the morphologically realized features are present also syntactically.

In Roberts & Holmberg (2010), the specific structural description of clitic incorporation and its morphological realization are not discussed in much detail. Assuming the valuation of the T features, as described in Sections 2.1 and 2.2, it appears that in Czech the valuation of T's [$u\varphi$] and [uD] features, which takes place in a syntactic relation of T with some overt DP, is signalled by two kinds of overt inflectional morphology:

(21)   a.   agreement morphology on the relevant parts of the predicate, as in (7), and
       b.   a nominative case marking on the overt subject DP.

I will not discuss here the nominative case marking of subjects, concentrating rather on the predicate's finite inflection, which reflects the valuation of the [φ]/[D] features.

3.1 Morphology and word order

As for Czech inflections, there is no argument at all suggesting that the Czech predicate (or DP subject) have to change their positions during the pre-Spell-Out derivation in order to enter an agreement relation signalled by their inflections. The finite verbal morphology itself does not require any movement of the v/V to

---

23 For more technical descriptions of more general processes of feature movement in the domain of the Probe, one can refer to Pesetsky (2013) and his process of Undermerge for nominal complexes reflected as a morphological case, or more generally, to Emonds (1987) and his process of morpheme insertion, labelled as Alternative Realisation. All the above processes can be described as a dislocation of syntactic (morpho-syntactic) features and, contrary to Baker (1985, 1988), if they are signalled by morphology, the morphemes surface to the right of the stem as in (20b). The domain of such right-side feature dislocation is very limited, usually to the first c-commanding or c-commanded lexical head.

the functional head T, there is no question inversion in Czech, and the PF position of the nominal complexes (including subjects) is not dependent on any version of case marking requirements, either. This claim may run counter to assumptions based on Baker's (1985) Mirror Principle and an incorporation approach to inflectional morphology. Nonetheless, it is justified by the Czech data: as it has been shown above, there is no clear word order restriction on the position of the Czech finite verb in the clauses with overt or null subjects; none of them requires overt dislocation of the constituents triggering or carrying the morphology listed in (21).

If we assume the GB theory of movement, the lack of movement in Czech may be explained by the Procrastinate Principle with morphology signalling a kind of post-Spell-Out movement at LF.[24] Within the minimalist copy theory of movement, one may argue that it is not always the highest copy that gets phonetically realized. I am not going to argue in favour of any particular explanation here, but I will take Czech inflectional morphology as a sufficient signal of a syntactic process, even when it is not accompanied by any overt movement of the inflected lexical entries.[25]

Having no clear argument in Czech based exclusively on the distribution of inflected lexical entries, one can wonder whether the distribution of features is indeed as suggested in (19). Are the features allowing the NS really located in some high verbal functional head T and not elsewhere? Does the checking process involve the subject in the domain of T at all? In the following sections, I provide some arguments in favour of T being the locus of the relevant features in Czech. First, however, I will describe in more detail the features present in the Czech agreement inflections, concentrating on their distribution in more complex predicate structures.

3.2 The agreement features in Czech

The mainly descriptive traditional grammar manuals of Czech (e.g., Grepl & Karlík 1998) state that its subject-predicate agreement reflects three nominal features: person, number and gender. A natural presumption then would be that these

---

24 Alternatively, we can assume that the syntactic movement does take place and it is obscured by some subsequent movement(s). This analysis, however, would require a lot of subsequent (and not clearly motivated) movements, especially if rightward dislocation were not available.

25 Discussing the data from a free word order language like Czech, one may wish to see a designated rightmost position (in both root and embedded clauses) as a landing site for constituents (verb included) discourse-marked for new/rhematic/focus interpretation. As long as I am not aware of any theoretically acceptable rightward dislocations, many Czech structures with clause final subjects have to be analysed as a result of many otherwise unmotivated internal merges (instances of left dislocation).

three features represent the [φ]/[D] feature content of the functional head T in (19a). The above three features, however, are not overtly present with every Czech finite predicate. In fact, they do not appear together in *any* single lexical entry, as the reader may have noticed on the basis of the examples above. This point will also be demonstrated in the following paragraphs in a more systematic manner.

First, let us consider the finite predicates that consist of one-word synthetic verbal form. These are considered the norm in the Czech present and synthetic (perfective) future tenses. The overt features in these paradigms are only the person and number features; we have already seen this verbal inflection in (7). To illustrate the paradigms, some randomly chosen examples are provided again in (22). In (22a), the 1ˢᵗ and the 2ⁿᵈ person forms are provided in both singular and plural, and (22b) illustrates the 3ʳᵈ person forms. Notice that none of these forms overtly encodes gender and still all these structures allow both overt and null subjects.

(22) a. *Najd-u /   musí-š /   udělá-me /   přeskočí-te*
        find-1SG must-2SG will.do-1PL will.jump.over-2PL
        'I find' / 'you must' / 'we will do' / 'you will jump over'

    b. (*On/ona/ono*) *najd-e /   mus-í /   uděl-ají /   přeskoč-í*
        he/she/it   find-3SG must-3SG will.do-3PL will.jump.over-3PL
        'he/she/it finds' / 'he/she/it must' / 'they will do' / 'they will jump over'

The Czech analytic future is illustrated in (23). The predicate consists of an infinitive of the lexical verb preceded by a future auxiliary (not a clitic) which spells out the same person and number features as the synthetic predicate in (22). In (23), we can compare the future auxiliary *bud-u* 'will.1SG' with other Czech lexical verbs subcategorizing for an infinitive ('can', 'start', 'want'). The infinitive following the future auxiliary is never inflected, i.e. it exhibits no φ-features.

(23) (*Já*)   *bud-u /   moh-u /   začn-u   pracovat.*
     I.NOM will-1SG can-1SG start-1SG work.INF
     'I will / can / start / (to) work.'

(24) (*Ty*)   *bud-eš /   můž-eš /   začn-eš   pracovat.*
     you.NOM will-2SG can-2SG start-2SG work.INF
     'You will / can / start / (to) work.'

(25) a. (*On/Ona/Ono*) *bud-e /   můž-e /   začn-e   pracovat.*
        he/she/it   will-3SG can-3SG / start-3SG work.INF
        'He/She/It will / can / start / (to) work.'

b. (*On/Ona/Ono*)   *bud-e* /   *mus-í* /   *začn-e* /   *chc-e*   *chodit.*
he/she/it   will-3SG   must-3SG   start-3SG   want-3SG   go.INF
'He/She/It will / must / start / want (to) go.'

Given that all the above verbal predicates allow NS, it appears that the Czech NSs are satisfied by the morphological agreement distinguishing overtly only person and number features. In terms of (19a), morphologically realized person and number features are sufficient to signal the checking of [$u\varphi$] and [uD]. If a gender feature is present in (22)–(25), it is not at all reflected in morphology.[26]

    Evidence for an agreeing gender feature can be found only in the analytic past and conditional predicates which combine the auxiliaries (some mentioned in Sections 3.3 and 3.4) with the *l*- and *n*-participles. First let us look at the Czech analytic preterit in (26a): the auxiliary has the form of a present tense verb *be*, while in (26b) the auxiliary has a special conditional form (for simplicity, I gloss it as *would*).

(26)   a.   (*Ty*)   *j-si*   *přišl-a.*
           you   be-2SG   arrive.PTCP-SG-F
           'You arrived.' (the addressee is a female)
     b.   (*Petr*)   *by*   *přišel.*
           Petr   would.3SG   arrive.PTCP.SG.M
           'Peter would arrive.'

In (26), the gloss indicates that the Czech *l*-participle (the active past participle) is inflected for number and gender features, but not for person. Moreover, we can see that the participles reflect the gender feature also with the 2nd person pronouns, which are not gender-marked in Czech.

    The same distribution is typical of the analytic passive, where the auxiliary (again, not a clitic) reflects only person and number features, while the passive *n*-participle reflects gender and number features. In (27), I use the 1st person subject pronoun, which (like the 2nd person pronoun) is genderless.

(27)   a.   (*Já*)   *j-sem*   *připraven-a.*
           I   be-1SG   prepared.PTCP-SG.F
           'I am prepared.' (speaker is a female)

---

26 A very detailed traditional descriptive summary of Czech agreements together with the taxonomy (in Czech) can be found in the New Encyclopaedia of Czech, see https://www.czechency.org/slovnik/SHODA .

b. (*Petr*) *je*   *připraven.*
Peter   be.3SG   prepared.PTCP.SG.M
'Peter is prepared.'

Based on Toman (1980) and Geist (1997), Veselovská (2003, 2004) demonstrates in detail that the partially homophonous past and passive auxiliaries show several morphosyntactic distinctions that demonstrate their distinct syntactic distributions. These distinctions, however, are irrelevant for describing the distribution of φ-features in Czech analytic predicates. Any initial part of the predicate in Czech (in either auxiliary or synthetic verbal forms) reflects the person and number features of the subject and none reflects gender.[27]

In structural terms, the label "initial" denotes the highest lexically realized part of the predicate, correlating with the position closest to the assumed subject position in Spec,TP. In this sense, the Czech person morphology is comparable with the English person agreement *-s* in the 3rd person singular (present tense). In the sections to follow, the position of agreement will be discussed in more detail.

3.3 Features in the T-domain

Contrary to most analyses which assume that the distribution of verbal clitics employs the mechanism responsible for the pronominal clitics, Kučerová (2014) proposes that neither of the Czech clitic auxiliaries moves to so-called C2 (clitic second) positions from some other lower verbal position. Rather, they are generated in the place they surface, i.e. in a high functional head in the T-domain. Most visibly, this claim is supported by the distribution of the negative prefix, which cannot appear on any clitic auxiliary; in the analytic verbal forms it must be prefixed to the following verbal participles. Accepting this proposal, in terms of the current analysis, the Czech clitic auxiliaries represent the morphology realized on T. In terms of (19), they signal the checked [uφ] and [uD] features on T. The T position of these agreement features is also supported by the following paradigms demonstrating the distribution of individual parts of the complex predicates containing clitic auxiliaries.

---

27 The only counter-examples are the 3rd person forms of the preterit, which contain no overt auxiliary. The participle may then seem to be the initial part of the predicate. This participle does not contain any overt person morphology (the 3rd person interpretation is based on the absence of the auxiliary), but it contains overt gender and number exponents. Therefore, also in this case, the uniformity of the paradigm forces us to assume the structural presence of a null 3rd person auxiliary (marked for person and number).

A functional sequence suggested by the Czech analytic verbal forms is presented in (26) and also in (28) below. The conditional clitic auxiliary *bych*, bolded in (28), is followed in an unmarked context by a past participle, which carries the number/gender morphology. The examples in (28) shows that a clitic auxiliary can be preceded by either one phrasal constituent: in (28a) it is a PP adverbial, and in (28b) it is the participle *přišla*.[28]

(28) a. *V   pondělí   **bych**   přišla.*
            on   Monday   would.1SG   come.PTCP.SG.F
            'On Monday I would come.'

      b. *Přišla   **bych**   dnes.*
            come.PTCP.SG.F   would.1SG   today
            'I would come today.'

The same auxiliary *bych* is used in (29) in the past conditional, where it is (in the standard paradigmatic ordering) followed by one or two other auxiliaries. The non-initial auxiliaries are not clitics but they appear in the form of an *l*-participle; they carry the same morphology as the following participle of the lexical verb (i.e., inflection *-a* showing the number and gender agreement). In (28b), we see that the participle of the lexical verb can be fronted, if (in a given paradigm) it immediately follows the auxiliary. Furthermore, we can see in (30a) that the first non-clitic auxiliary can be fronted to precede the clitic. However, the verbal elements that do not immediately follow the auxiliary cannot be fronted over it, as shown in (30b-c); more than one verbal element cannot be fronted either, (30d).

(29) *Minulé   pondělí   **bych**   byla   (bývala)   přišla*
        last   Monday   would.1SG   be.PTCP.SG.F   be.PTCP.SG.F   come.PTCP.SG.F
        (i   já).
        even   I.NOM
        'Last Monday (even) I would have arrived on time.'

(30) a. *Byla*   **bych** t_{byla} *bývala   přišla   minulé pondělí   (i ja).*
      b. * *Bývala*   **bych**   *byla* $t_{bývala}$ *přišla   minulé pondělí   (i ja).*
      c. * *Přišla*   **bych**   *byla   bývala* $t_{přišla}$ *minulé pondělí   (i ja).*
      d. * *Byla bývala přišla*   **bych** t_{byla} $t_{bývala}$ $t_{přišla}$ *minulé pondělí   (i ja).*

---

28 Czech C2 clitics follow an XP (phrase) or V/C (head). The exceptions include examples when the initial constituent remains covert, or the weak complementizer does not count as first. For more discussion see e.g. Toman (1999), Veselovská (1994), Cardinaletti & Starke (1999), Franks & King (2000), Rezac (2005) and Dotlačil (2007).

The adjacency requirement on the fronted verbal element and the contrast between (28b)/(30a) and (30b, c) argue in favour of a structure with a sequence of individual verbal (auxiliary) functional heads above a lexical verb. In this functional sequence, the position of auxiliaries is fixed and the top position is occupied by the clitic auxiliary. On the other hand, the lexical verb is the bottom of the verbal projection. Notice that with the exception of the clitic, all other verbal elements show gender and number (but not person) agreement.

The high position of the person agreement features in Czech is also shown by the existence of inflected forms of some complementizers that contain the conditional auxiliary *by* (and its agreeing φ-features). In (31a) we can see the resultative *aby* 'so that' and in (31b) the conditional *kdyby* 'if'. Both of these complementizers are used in embedded clauses together with *l*-participles and both transparently contain the morpheme identical with the conditional auxiliary *by* 'would', in (28) and (29).

(31) a. *Chtěl,*   **abych**   (*já*)   *odešel*   (*já*)
       want.PTCP.(3)SG.M so.that.1SG I.NOM leave.PTCP.SG.M I.NOM
       'He wanted me (a male) to leave.'
    b. *Bála*   *by*   *se,* *kdy***bych** (*já*)   *odešel*
       fear.PTCP.SG.F would.3SG RFL   if.2SG   I.NOM leave.PTCP.SG.M
       (*já*)
       I.NOM
       'She would be scared, if I (a male) left.'

The position of inflected complementizers in (31) must be at the very top of the embedded structure (immediately preceding the clitic cluster), and their morphology represents an incorporation of subject features to C. The fact that such incorporation can take place suggests that the subject features are indeed at the relevant phase of derivation very close to C. According to the analysis proposed here they are in T because of the preceding D-to-T clitic incorporation.

All the examples in (28), (29) and (31) can, but need not, contain overt pronominal subjects. The fact that the pronominal subjects can be located at the very end may suggest that the lexical verb is not inside the vP. However, at the same time, the presence of (several) auxiliaries makes it clear that the lexical verb position is not in T, either. Crucially, the examples (28), (29) and (31) demonstrate that (i) the source of the agreeing person and number features is very high in the functional domain, plausibly in T, and (ii) there is another field of agreement (for gender and number features) in the lower verbal domain.

Moreover, the morphology reflecting the valuation of T features (i.e. person and number features) can get realized in several positions in Czech, i.e. on

inflected complementizers above T, on clitic auxiliaries in T, and in the first overt verbal head below T.[29] This kind of distribution is standard for strictly local inflectional morphology and can be described in terms of e.g. incorporation, affix hopping, feature fission or alternative realization – depending on the specific framework. Descriptively, Veselovská (1994) already demonstrated that although the position of the Czech finite (i.e. inflected) verb is considerably free with respect to its phrasal constituents, there is never any overt functional head between a theoretically assumed functional head T and the verbal form inflected for person and number agreement (in Czech, a finite verb and the functional head T above it are arguably in the same local domain).

3.4 Rich versus poor morphology

We have seen in all the examples in the preceding sections that there is no distinction between the Czech verbal paradigms used with overt subjects and with null subjects. This suggests that the so-called "rich inflection" that licenses NS (indicating D-to-T clitic incorporation) is phonetically identical with the inflection reflecting the feature checking in the presence of an overt subject DP. To keep the difference in mind, let us call the latter "poor inflection". Even though the processes leading to the agreement morphology may be structurally distinct, as suggested by the two diagrams in (20), both processes are signalled by a single set of syncretic morphemes (listed in (7)).

If we look more carefully for a distinction between rich and pure inflection, there is one paradigm suggesting that such a distinction really exists and can be discovered in colloquial Czech.

Below I illustrate the colloquial paradigm of the Czech analytic preterit. First, look at (32), which shows the standard analytic verbal form of the 1st person singular (clitic) auxiliary *jsem* 'am', followed by a past participle marked for gender and number. In (32a), the unmarked linear order illustrates an overt pronominal subject preceding the auxiliary clitic. Because it is the host of the clitic, it cannot be dropped unless replaced by another host constituent (e.g., an adverbial). In (32b), we see that a past participle can also precede the clitic (as is standard with the C2 clitics), and that an overt pronominal subject can appear elsewhere, preferably at the very end, due to the fact that an overt pronoun will be discourse-marked

---

29 In analytic future it is a future auxiliary (see (23–25)), in synthetic forms it is a synthetic predicate (see (22)).

as a focus. Unless it supports the clitic, the non-focused pronominal subject can always be dropped.

(32) a. *Já* / *pro   *včera* /   *jsem*   *mu*   *to*   *nedala.*
     I.NOM /(I.NOM)   yesterday   be.1SG   him   it   not.give.PTCP.SG.F
     'I did not give it to him yesterday.'

     b. *Nedala*   *jsem*   *mu*   *to*   (?*já*)   *včera*   (*já*).
       not.give.PTCP.SG.F   be.1SG   him.DAT   it.ACC   I.NOM   yesterday   I.NOM
       'I did not give it to him yesterday.'

I assume that in the absence of an overt subject, the auxiliary *jsem* in (32) represents the agreement which results from the D-to-T clitic incorporation, i.e. rich agreement (see the discussion at the beginning of Section 4). When the pronominal subject is present, the auxiliary can be taken as a verb realizing only pure agreement features. Notice again that in both cases the agreement is realized by the same morpheme and so in both cases is located high in the verbal domain (the auxiliary is a clitic with both overt and null subjects).

     In colloquial Czech, the 1st person auxiliary *jsem* 'am' can be omitted in the presence of the overt subject pronoun *já* 'I', as illustrated in (33).[30]

(33) ***Já***   (*jsem*)   *mu*   *to*   *nedala.*
     I.NOM   be.1SG   him.DAT   it.ACC   not.gave.PTCP.SG.F
     'I did not give it to him.'

The acceptability of auxiliary deletion shows that in 1st person past, in colloquial Czech, there are two ways to economize on the analytic predicate: one is to keep the auxiliary and to use the NS – the structure shared with the standard dialect, as in (32b). Alternatively, colloquial Czech can also keep the overt subject pronoun *já*, as in (33), and drop the agreeing auxiliary *jsem*, which realizes the same φ-features (person and number).

     This paradigm is interesting for yet another reason. In (32) above we have seen that the full pronominal subject *já* 'I', co-occurring with the past auxiliary, can appear in any order with respect to the auxiliary – initial, middle, or final position. However, the same is not true with the pronominal subject *já* in the absence of the past auxiliary. The only fully acceptable structure is (33), where the pronoun is clause-initial. The contrasted, unacceptable orderings follow

---

30 The Czech past (clitic) auxiliary deletion in the presence of the overt pronoun is restricted to the 1st person. In the singular (illustrated here) it is quite frequent, while in the plural it is rarer.

below. In (34a), the auxiliary-substitute *já* 'I' follows the participle and, in (34b), it follows the adverb – both of which are acceptable in (32).[31]

(34) a. *Nedala*   (*_já_)   *mu*   *to*   (*_já_).
       not.gave.PTCP.SG.F I.NOM   him.DAT   it.ACC   I.NOM
       'I did not give it to him.'

    b. *Včera*   (*_já_)   *mu*   *to*   (*_já_)   *nedala*   (*_já_)
       yesterday I.NOM   him.DAT   it.ACC   I.NOM   not.gave.PTCP.SG.F   I.NOM
       'I did not give it to him (yesterday).'

It is obvious that we cannot call the pronoun *já* in (33) a clitic, because it does not take a clitic position (recall that Czech clitics cannot be clause-initial and they must follow the first XP/V constituent). Neither of these criteria holds for the overt *já* substituting for the auxiliary. Thus, if the auxiliary-substitute *já* is not a clitic, the ungrammaticality of (34a, b) could be due to the C2 pronominal clitic sequence *mu to* 'him it'.

However, the same explanation is not available for the position after the clitic cluster and the clause-final position in (34b), which is very significant because the overt pronoun *já* 'I' is always interpreted as a part of the rhematic (focus, new) information, while clitics are prototypical themes (old information elements). The unmarked position of the focus is at the end of the sentence in Czech, and the initial position carries the focus feature only in combination with the contrastive stress.

Therefore, I propose to relate the ungrammaticality of overt subjects in (34) to the distinction between the rich and poor inflectional morphology.

The auxiliary combined with the NS, as in (32b) or (33), represents the rich morphology, corresponding to (20b). On the other hand, the morphology on the auxiliary co-occurring with the overt pronoun in (32) represents the poor agreement inflection resulting from the structure in (20a). I propose that colloquial Czech, in fact, avoids the doubling of the overt subject pronoun features with their identical marking in the pure agreement pattern.

The ungrammaticality of overt subjects in (34) is also relevant with respect to the position of the overt subject. The examples suggest that in the absence of the rich agreement (i.e., D-to-T incorporation), the unvalued features of T have to be checked via the overt movement of the pronominal subject to T, as in (20a). In

---

31 The sentences in (34) are acceptable with the 3rd person singular NS, because in the preterit paradigm the 3rd person auxiliary is null and the participle shows a singular agreement. See also footnote 28.

other words, in the absence of (20b), i.e. the absence of the D-to-T incorporation signalled by the person and number agreement, the overt structure (20a) with the subject in Spec,TP is enforced. And because the 1ˢᵗ person pronoun is not a clitic in Czech, the only available position for this pronoun is clause-initial, preceding the clitic cluster.[32] The availability of post-verbal subjects in NSLs, as stated in the parametric cluster in (2), can be explained using the same principle: the post-verbal subjects signal the availability of the D-to-T incorporation, as depicted in (20b). If this process is not available in a given language, (20a) may become (the only?) option. I believe that there is a potentially interesting direction for future research here, which will also attempt to explain why (or predict that) rich agreement and pure inflections are frequently syncretic.

4 Two feature sets

In the preceding section, I concluded that Czech verbal agreement (in person and number) is related to the feature content of the functional head T. I also demonstrated that the agreeing φ-features in T can be accompanied by the gender and number features realized in the lower verbal domain. In this section, I repeat some of the data already mentioned in the sections above and provide additional data in order to sketch a more precise picture of the distribution of these two feature sets.

In section 3, we saw that the feature complexes reflected in agreement in Czech analytic verbal predicates are very similar. In all cases, the initial lexical form, which is located in some verbal functional head, contains person and number features. With synthetic predicates, this represents the whole agreement inflection, with both overt and null subjects. With analytic verbal predicates (the non-initial agreeing elements) take the form of participles and they show the number and gender features, but not the person feature.[33]

The Czech preterit and passive participles are discussed in Veselovská (2004) and also in more detail in Veselovská & Karlík (2004). The authors propose that

---

32 Why is it restricted to 1ˢᵗ person only? I propose it may have something in common with the unmarked setting of the 1ˢᵗ person in the participant role of the discourse couple (speaker vs. hearer/addressee).

33 As for the terminology of feature distribution and the categories of lexical stems, the text here will use the label T(P) for the functional head above vP, signalled by the person and number features, and the label vP for the position of the participles (the categorized lexical stem of the verbal predicate) marked by the gender and number features.

both the *l*-participle in preterits and the *n*-participle in passives appear in a low functional verbal projection. Moreover, their analysis demonstrates that both Czech participles show a very restricted set of verbal characteristics and a correspondingly augmented set of properties typical of adjectives. Therefore, Veselovská & Karlík (2004) conclude that Czech participles are best analysed as PF-derived adjectives (APs).³⁴

On the other hand, Veselovská & Karlík (2004) point out that the feature set including the person feature is special, because – contrary to the set including the gender feature – it is unique. It is the only part of the predicate that reflects person specification, and, considering linearity, it is invariably the first (hierarchically the highest) part of the analytic verbal predicate where person combines with number into a kind of fused suffix. Because of their high (initial) position, person (and number) features are the best candidates for being the Probe features in T, stated as [*u*φ] and [*u*D] in (19a). The uninterpretable Probe features delete after being checked (which often makes them morphologically visible).

With regard to gender (on participles), this feature can appear in Czech on analytic predicates several times. In (35), I illustrate the Czech past conditional passive with four auxiliary (functional) verbs: one clitic auxiliary *bys* and three *l*-participles: the first *byla* is a part of the analytic past, the second *bývala* marks aspect, and the third *byla* is part of analytic passive. At the very end, the passive *n*-participle *představena* contains the lexical root. Notice that only the initial (finite) auxiliary *bys* shows person and number features. On the other hand, all the participles identically reflect the same gender (and number) features.

(35) (*Ty*)   *by-**s***   *byl-**a***   *býval-**a***   *byl-**a***
       you   COND-2SG   be.PTCP -SG.F   be.PTCP- SG.F   be.PTCP-SG.F
       *představen-**a***.
       introduce.PTCP-SG.F
       'You would have been introduced.'

The example in (36) demonstrates that, apart from the main predicate, other parts of the clause also show agreement with the subject nominal. Apart from the past participle of the main verb *odešla* 'leave.SG.F', the secondary predicates

---

34 To describe the adjectival characteristics of Czech participles, Veselovská & Karlík 2004 apply both the semantic and the morpho-syntactic diagnostics proposed in Wasow (1977). For the Slavic data, such tests were already used by Levin & Rappaport (1989) and Schoorlemmer (1995). For argumentation concerning the PF status of adjectival inflection (including in Czech), see Emonds (2012).

sama 'alone.SG.F' and *spokojená* 'satisfied.SG.F', and the transgressive *usmívajíc se* 'smiling RFL' are recursive, showing gender and number features.

(36) (*Jana*)   *by*   *odešl-**a***   *sam-**a***   *ale*
     Jane.NOM   would.3SG   leave.PTCP -SG.F   alone-SG.F   but
     *spokojen-**á***
     satisfied-SG.F.NOM
     a   *usmívaj-**íc***   *se*
     and   smiling-SG.F.NOM   RFL
     'Jane would leave alone but satisfied and smiling.'

As we can see, the secondary predicates and the transgressive in (36) do not reflect person features. If the subject is changed to the 1st person, only the first auxiliary will reflect the change; if the speaker is a female, the agreements in the secondary predicates and the transgressives remain the same (compare (36) with (37)).

(37) (*Já*)   *bych*   *odešl-**a***   *sam-**a***   *ale spokojen-**á***
     I.NOM   would.1SG   leave.PTCP-SG.F   alone-SG.F   but   satisfied-SG.F.NOM
     a   *usmívaj-**íc***   se
     and   smiling-SG.F.NOM   RFL
     'I (a female) would leave alone but satisfied and smiling.'

The distinct characteristics of the auxiliary φ-feature set (person and number) and the secondary predicate / transgressive feature set are further evidenced in examples (38) and (39). They show that the latter agreement is not restricted to a structural case position. The same agreements are triggered by a structurally accusative object, as in (38), and by an oblique locative case object of a preposition, as in (39) (the adjectival agreement also includes the morphological case feature, accusative or locative, required by v/V or P, respectively).[35]

(38) *Viděla*   *by*   *Petra*   *spokojen-**ého***   *a*
     see.PTCP.SG.F   COND.3SG   Peter.SG.M.ACC   satisfied-SG.M.ACC   and
     *usmívaj-**ícího***   *se.*
     smiling-SG.M.ACC   RFL
     'She would see Peter satisfied and smiling.'

---

35 Based on the analogy with case marking in (38), where the adjectivals agree in case, Veselovská & Karlík (2004) propose that the participial adjectival suffixes with both *l*- and *n*-participles realize the (default) nominative case.

(39) *Mluvila*   *by*   *o*   *Petrovi*   *spokojen-**ém***
      speak.PTCP.SG.F COND.3SG about Peter.SG.M.LOC satisfied-SG.M.LOC
      a   *usmívaj-**ícím***   SE
      and   smiling-SG.M.LOC   RFL
      'She would speak about Peter (as) satisfied and smiling.'

I am not going to argue here for any specific structure underlying the secondary predicate and the transgressive agreements. However, I would like to argue at this point that this agreement is distinct from the one in the initial parts of the predicates (in the T-domain). In (40), we can see that the nominal complexes agreeing with the V or P objects cannot be dropped in the same way as the subject in (36). That is, in Czech there are no null objects similar to NS, which would be licensed by the agreement on secondary predicates or transgressives.

(40) a. * *Jana*   *viděla*   pro   *spokojen-**ého***   *a*
             Jana.NOM see.PTCP.(3)SG.F (3SG) satisfied-SG.M.ACC and
             *usmívaj-**ícího***   *se.*
             smiling-SG.M.ACC   RFL
             'Jana saw (someone) satisfied but smiling.'
      b. * *Jana*   *mluvila*   *o*   pro   *spokojen-**ém***
             Jana.NOM speak.PTCP (3)SG.F about (3SG) satisfied-SG.M.LOC
             a   *usmívaj-**ícím***   se
             and   smiling-SG.M.LOC   RFL
             'Jane spoke about someone satisfied and smiling.'

On the other hand, the elliptical omission of a noun inside a complex DP which contains a rich internal agreement is otherwise acceptable in Czech, especially when such a DP contains a demonstrative. Example (41) shows that the features corresponding to the English pronoun *one* are represented by agreement on adjectives (and demonstratives) in Czech.

(41) *Vezmi*   *si*   *tu*   *novou*   [$_N$ __].
      take.2SG.IMP RFL the.ACC new.ACC (one)
      'Take the new one.'

The existence of a covert N in structures like (40) is not a property of the NS parametric cluster. On the other hand, the characteristics of the consistent NSL in (19) include a requirement that a NSL have DPs where the n/NP complement serves as a variable bound by $D_{[\varphi]}$. Given (19b), I assume that the Czech DP can contain a covert *pro* in the n/NP domain: [$_{DP}$ $D_{[\varphi]}$. . . [$_{n/NP}$ *pro*]].

4.1 Index, concord and feature economy

The distinct nature of the person (and number) and gender (and number) agreement feature sets, illustrated in the preceding section, is foreshadowed in traditional terminology. When describing the typologically distinct morpho-syntactic kinds of feature sharing, traditional grammarians distinguish two kinds of agreement one of which is comprised of person and the other of gender, as in (42).[36]

(42)   a.   INDEX (the verbal agreement) reflects person and number features.
       b.   CONCORD (the adjectival agreement) reflects gender and number features.

Notice that the taxonomic statement presented in (42) refers to the features of person, number and gender, while (19**)** mentions D- and φ-features. Looking for correlates, the feature of person, typical of pronominals, is plausibly a part of the feature set of category D in (19).

The status of the number feature is less clear, since it appears in both (42a) and (42b): both the initial auxiliaries and the following participles and other adjectival agreements reflect the number feature. Given its presence with auxiliaries, this feature has to be (one of) the uninterpretable (strong) features of the Probe in T, i.e. the feature requiring a unique checking operation (elimination). The person and number features are then arguably the features of the incorporated clitic licensing NS. As for the position of the number feature inside the Goal (the DP projection), it is quite uncontroversial that it is present in the functional sequence in the extended nominal projection, as Q or Num. I will not distinguish Q/Num and D here, assuming that person and number are both in the high functional domain of the DP.

On the other hand, the number feature is also part of the normal Czech adjectival concord in (42b). With category A, the number feature can be valued recursively in the specific structural contexts, including participial agreement. The number feature shares then the characteristics of the gender feature.

Considering this feature in Czech, along with [+human], gender denotes a referent's natural sex and therefore qualifies as an intrinsic interpretable φ-feature of a noun. If gender realizes some specific functional head, this head is very low and close to a lexical category. And gender in Czech, as seen above, is never part of the synthetic predicate inflection that serves to license NS.

---

36 A detailed discussion in the context of current typological frameworks (including Slavic) can be found in e.g., Corbett (2006, 2012) and Wechsler & Zlatić (2003).

We can therefore argue that gender, though present syntactically as one of T's φ-features, is not overtly realized in Czech finite morphology, due to some kind of morphological economy. In Distributed Morphology (e.g., Noyer 1998; Nevins & Parrott 2010 and Parrott 2015), feature economy is redefined as a process of impoverishment, which can be used to explain the lack of morphological realization of several marked pronominal features. Another possibility is a general principle of feature economy proposed in Emonds (2012) to explain the presence/absence of overtly realized feminine forms in Czech adjectival morphology. This author claims that generally (cross-linguistically) no attested inflectional morphology ever fuses more than two marked morphemes, one of which is always a number exponent.

On the other hand, the person and number features realized uniquely on the initial finite predicate is identical with the features of the genderless 1st and 2nd person pronouns. This identity of feature sets supports an analysis which assumes that verbal agreement represents some kind of incorporation of a unique clitic pronoun. The feature economy then plausibly refers to the repertory of pronouns, and verbal agreement simply reflects it.

It may also be that the gender features on participles (i.e., the concord features) are not in fact relevant for the probing features in T, namely [$u\varphi$] and [uD]**.** Instead, gender features may take part in the structure-binding process unifying the feature content of the DP. Such unification of feature sets may be part of the process of providing reference to the subject, mentioned in (19b) as a stage required to license the NS in a consistent NSL. However, I will not pursue this analysis further at this point.

Instead, I would like to provide more arguments in favour of the claim that the traditionally single concept of subject-predicate agreement licensing NS in fact involves two independent (separate) feature sets. In the next section, I will show that the two sets of features are checked separately in at least two distinct domains. This argumentation further supports the idea that number features mentioned twice in (42) – as being part of both agreement and concord – refer in fact to two distinct feature sets.

4.2 Feature conflicts

In numerous linguistic studies, it is observed that apart from standard agreement patterns based on some language-specific rules, there are also cases which contradict the normal rules. Given the lists of those apparent violations of standard agreements, it appears that the contexts in which such morphosyntactic structures appear are cross-linguistically very similar. Czech represents such a

language: it displays most of these contexts for "disagreement" (or feature mis-matches).[37] Those listed in (43) are divided according to the characteristics of the subjects entering into agreement.[38]

(43)   The subject of the subject-predicate disagreement structure involves
    a.   a special lexical item
    b.   a coordinated constituent
    c.   a quantified noun phrase
    d.   certain kinds of pronoun subjects

As for (43a), in Czech this kind of agreement violation is restricted to only one lexical entry: *děvče* 'girl', which denotes a young female, but systematically triggers the neuter gender in both index and concord agreements, plausibly because all young animates (e.g., the lexical entries for children as well as for kittens or goslings) are marked with neuter in Czech.

The agreement with coordinated structures in (43b) is widely discussed in the generative framework, especially because it is the only context in Czech in which pre-verbal and post-verbal subjects result in distinct agreements. Skrabalova (2007, 2008) provides analyses explaining why the pre-verbal coordinates usually trigger plural agreement, while post-verbal subjects also allow agreement with the first conjunct. The feature symmetry of coordinates is also discussed in Kučerová (2017), who points out some unexpected patterns occurring in quantified coordinates and neuter coordinates. Since I cannot see any interesting relation to the topic discussed here, I will not describe those variations here in more detail. For the same reason, I will not demonstrate or analyze examples containing quantified DPs mentioned in (43c), either.[39]

The examples relevant for discussing how distinct sets of features participate in subject-predicate agreement (and NS licensing) concern agreements with pronouns, listed as (43d).

---

37 Some of the data and many more similar examples for all the structures demonstrated here are provided in Veselovská (2003, 2018, 2020), with analyses compatible with the one provided here.
38 For more Czech data, see Panevová & Petkevič (1997) and Kučerová (2000). Similar examples in typologically more varied languages can also be found (Corbett 2006, 2012).
39 In most Slavic languages, agreement depends on the kind of quantifier. The morphology attested in these complex and theoretically challenging structures provides insight into the DP projections, including a hierarchically ordered taxonomy of φ-features. For an extensive discussion and a range of analyses reflecting the stages of generative approaches, see Babby (1987), Pesetsky (1982, 2013) and Franks (1995), among others. For Czech, see Veselovská (2001, 2018).

First, we have already seen in (26)–(35) and (37) that the 1st and 2nd person Czech pronouns are not gender-marked. Still, as seen below, they trigger concord agreement, including the gender feature on participles and elsewhere. Example (44) demonstrates that the subject *já* 'I' is not marked for gender and yet it can trigger both feminine and masculine agreement on verbal participles, secondary predicates and transgressives.

(44) (*Já*)   *bych*   *odešel-Ø/-a*   *spokojen-ý/-á*   *a*
     I.NOM   COND.1SG   leave.PTCP-SG.M/F   satisfied-SG.M/F   and
     usmívaj-e/-íc   *se.*
     smiling-SG.M/F   RFL
     'I (male/female would leave satisfied and smiling.'

This example is puzzling, but it can be explained by some sort of feature economy as discussed in the preceding section. Assuming that the speech act participants (i.e., the speaker and the addressee) are marked for values of person (both discourse participants), we may conclude that the gender feature must be syntactically present but morphologically unrealised on the overt 1st and 2nd person pronouns. Nor is gender overt on the finite (index) agreement involving the 1st and the 2nd person pronouns incorporated into the predicate.[40] In these cases, the structures with genderless pronouns do not give rise to any examples of a feature mismatch.

More plausible candidates for real "disagreement" within one verbal domain can be found with Czech polite pronominal subjects (or their NS variants). In Czech, the hearer (2nd person) can be addressed either with the singular *ty* 'you.2SG', in case of a typical equal or close relation between the speaker and the addressee, or with the plural *vy* 'you.2PL', in case the single hearer/addressee is hierarchically higher or more distant. Example (45) illustrates the plural pronoun *vy*: the pronominal feature set (person and number) is reflected in the auxiliary – the agreement morphology with *vy* is invariably 2nd person plural. The past participle, however, as well as a secondary predicate or a transgressive, is obligatorily marked for singular (the actual number of the addressee) and this number is distinct from the number of the auxiliary. The participles also reflect gender, masculine or feminine, depending on the actual gender of the addressee.

---

40 The structures with genderless pronouns are described in detail in Bonet (1991). For a more detailed discussion in the functionalist, HPSG and DM frameworks, see Veselovská (2018, chapters 8 and 9) and the literature cited there.

(45) (*Vy*)   *by-ste*   *odešel-Ø/-a*   *spokojen-ý/-á*   *a*
you.2PL.NOM COND-2PL leave.PTCP-SG.M/F satisfied-SG.M/F and
usmívaj-e/-íc   *se.*
smiling-SG.M/F RFL
'You would leave satisfied and smiling.' (the hearer, male or female, is hierarchically higher or more distant)

The example in (45) thus demonstrates a clear case of disagreement in number. Within one predicate we find conflicting features in the index and concord agreements; the first is in plural and the second is in singular. Traditionally, descriptive linguistic frameworks call the agreement reflected on the auxiliary in (45) a "formal agreement", because it reflects the plural form of the polite pronoun *vy* 'you'. The agreement on the participle, on the other hand, is called a "semantic agreement", because it reflects the fact that the addressee is a single person, male or female. The descriptive frameworks do not try to explain why one part of the predicate follows the formal agreement, while the other follows the semantic one.[41]

In the analysis pursued here, the auxiliary agreement reflects the incorporation of the polite pronoun features into T (the Probe) – therefore the features of the subject and the finite verb are identical. These features value features [uD] and [uφ] of the Probe. Moreover, these features plausibly reflect the Topic functional domain (related to the discourse) and not the base-generated position of the subject.

The feature set of the subject's base-generated position is expressed by the concord features spelled out on the participle – the semantic feature set, including gender and number values, are clearly distinct from the number feature of T (see the initial auxiliary).

Veselovská (2019, 2020) proposes that the auxiliary reflects the features of the functional domain of the subject DP (checked against the Topic projection), while the participle agrees with the features of a phonetically unrealized n/N that is nonetheless syntactically present inside the genderless DP pronoun. This covert n/NP contains the syntactic representation of the semantic (real) features

---

41 A functionalist discussion of the structures with polite pronouns can be found in Brown & Gillman (1960), who refer to plurality as a traditional metaphor for power or "social authority". In a more recent framework, a similar analysis is given in Wechsler & Zlatić (2003). These authors assume a wider repertory of pronouns with a wide variety of feature contents. None of the approaches explains the source of the conflict of features within one predicate. Perhaps some resolution rules for the index morphology in T could be used – because in the conflicting patterns, the index features always represent a more marked version when compared with the conflicting concord features, but this point will not be analysed here.

of the addressee. The disagreement structures like (45) demonstrate that Czech morphology does not conflate (unify or resolve) the potentially conflicting gender and number features of the lower lexical N-domain with the functional D-domain (containing person and number features) in any single inflectional morpheme.

Using the formalism in (19b), the feature distribution inside the structures with the polite pronoun *vy* can now be schematically represented as follows (the labels of the features are simplified):

(46)   *vy* 'you.2PL' = [$_{DP}$ [$_{Person}$ 2] [$_{Number}$ PL] . . . [$_{n/NP}$ [$_{Gender}$ F/M] [$_{Number}$ SG]]]

The location of features illustrated in (46) is also supported by the agreements triggered by Czech animate pronoun *někdo* 'somebody'. Based on word order and feature content (the presence of animacy and gender) and case agreement patterns, Veselovská (2003) argues that this animate indefinite pronoun *někdo* realizes a nominal n/N head which has a D-layer. In the following examples, this pronoun *někdo* with the 3ʳᵈ person singular feature content can in fact combine with a NS representing a plural polite subject *vy*. In (47), the agreement on the auxiliary reflects the plural feature of the NS, while the participle shows the gender and number features of the pronoun *někdo*, generated under n/N. Note that only masculine unmarked gender is available and, therefore, the term "semantic agreement" would not be justified in this case.

(47) a. pro   *někdo*   byste   **mohl-Ø/*-a**   pomoci.
          (2PL) somebody.3SG.M   COND.2PL   can.PTCP-SG.M/F   help.INF
          'Someone should be able to help.'
      b. pro   *někdo*   to   *už* konečně   **přines/-te.**
          (2PL) someone.3SG.M   it.ACC   at last   bring.IMP.2SG/-2PL
          'Someone, bring it at last.'

This structure can be both declarative, as in (47a), and imperative, as in (47b); it is the synthetic predicate verb *přines/-te* 'bring' that reflects the 2ⁿᵈ person singular or plural NS (plural allows the polite interpretation). The examples above confirm that the proposed separate and potentially distinct sets of features participating in subject-predicate agreement are indeed attested in the data. We have seen that the subject-predicate agreement licensing the NS is not a one-step syntactic operation, but a sequence of several steps.

5 Summary

Following a brief introduction to the *pro*-drop phenomena, this chapter has demonstrated in Section 1 that the Czech data robustly exhibit all the widely proposed characteristics of the *pro*-drop cluster, stated in (2) and repeated below for convenience.

(48) *Pro*-drop/NS-parametric cluster
- a. existence of *pro* – a silent null subject of a finite clause;
- b. no active *that*-trace effect;
- c. both pre- and post-verbal overt subjects possible (free subject inversion);
- d. rich (alternatively, "uniform") inflection across person and number specifications;
- e. a definite (or specific) interpretation of 3SG *pro*/NS (marked *pro*/NS$_{arb}$).

As we have seen, Czech finite predicates, although located in a low position both preceding and following overt subjects, have rich agreement morphology that allows for a NS. The NS carries a non-contrastive definite (specific) interpretation in all tense/aspect contexts for all three persons in both singular and plural. On the other hand, standard Czech generic subjects can be covert only with agreement distinct from 3rd person singular or in impersonal passives. This chapter has also provided some examples of idiomatic structures and constructions with (i) an obligatory NS and (ii) a precluded NS.

In Section 2 I argued that the Czech data can be described and analysed in the theoretical framework developed for NSLs by Holmberg (2005, 2010). This analysis assumes the minimalist process of checking the uninterpretable (strong) nominal features in the verbal functional head T, as stated in (19) and repeated here as (49).

(49) Features of T/D in a consistent NSL
- a. Probe: T [*u*φ], [*u*D]
- b. Goal: [$_{DP}$ D$_{[φ]}$... [$_{n/NP}$ *one*]] or [$_{DP}$ D$_{[φ]}$... [$_{n/NP}$ pro]]

In Section 3 I have attempted to correlate the φ-features needed for licensing NS with more specific traditional nominal categories. The φ-feature distribution in Czech indicates that there are two distinct feature sets, one of which is unique in the clause (the index agreement in person and number) and the other recursive (the concord agreement in gender and number). These two sets of features participate in the agreement process licensing NS in distinct ways (and phases).

Finally, Section 4 has provided support for the claim concerning the independence of the two feature sets providing examples of more complex Czech predicates in which the index agreement morphology conflicts with the concord agreement morphology. Based on the data the process of agreement was described, following the proposal made in (49).

According to (49a), the NS-licensing agreement concerns the valuation of the uninterpretable features of T. The feature content of T is valued with respect to a discourse functional head in an articulated C-domain. The Topic functional head plays a crucial role in the identification of subject and primarily allows the valuation of the person feature. In Czech, the number feature (the one present in the index agreement) is part of the verbal inflection – a portmanteau morpheme fusing the person and number features. The index agreement inflection signals a unique syntactic process and is reflected in Czech by the morphology on (i) inflected complementizers, (ii) the clitic auxiliaries, and (iii) synthetic finite verbs.

The above process can be described as a D-to-T pronominal clitic incorporation as long as the features realized in T are identical with the features of the relevant 1st and 2nd person subject pronouns (i.e., the φ/D-features). In (20b), the process of D-to-T incorporation is presented as the movement of features from the subject base position. If, instead, the φ/D-features of T were valued by the corresponding person and number features of the Topic head, the checking relationship might not need the features of the vP internal subject at all. The modified structure, including a Topic projection, is given in (50), where the φ/D-features of the Topic head incorporate into T.

(50)   TopicP
        /   \
    Topic   TP
   [Person]   / \
   [Number] T   vP
          [uφ][uD] / \
                      DP
                   [Gender]
                   [Number]
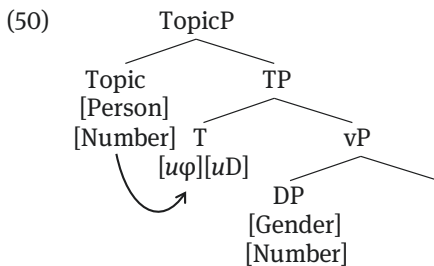
As for concord morphology, the gender and number features are interpretable features present in the lower verbal domain, plausibly in the subject base-generation position inside the vP. In Czech, these features can be recursively reflected on participles, secondary predicates, or transgressives. The Probe for these features is clearly not a verbal but an adjectival head, realized as a portmanteau morpheme fusing gender and number (as well as case) features. In the concord type of agreement, the person feature is irrelevant.

It remains to be determined whether the gender feature is part of the valuation of T's uninterpretable features or whether the features provided by the Topic head (person and number) are indeed the only source of the φ-set incorporated into T. If the former scenario were right, the valuation process of T's [*u*φ] and [*u*D] would target two Goals – the features of the Topic head and the features of the subject – with some features overtly realized in the inflection and others kept unrealized due to feature economy. In that case, however, one would not expect feature conflicts described in 4.2. On the other hand, it is very plausible that the concord gender and number features play a role in the binding of the φ-features inside the subject DP required in (49b). This process provides the subject with reference and ensures appropriate uniformity of N's and D's φ-feature sets (at least the uniformity of number). However, the data in this chapter demonstrate that this unification is not signalled overtly in the inflection morphology of Czech analytic predicates.

References

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing Agr: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539.

Alexiadou, Artemis. 2006. Uniform and non-uniform aspects of *pro*-drop languages. In Peter Ackema, Patrick Brandt, Maike Schoorlemmer & Fred Weerman (eds.), *Arguments and Agreement*, 127–158. Oxford: Oxford University Press.

Babby, Leonard H. 1987. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language and Linguistic Theory* 5. 91–138.

Baker, Mark C. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16 (3). 373–415.

Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Barbosa, Pilar. 2019. *Pro* as a minimal nP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50 (3). 487–526.

Bartošová, Jitka. 2017. *Topics in Copular Clauses*. PhD dissertation. McMaster University, Hamilton, ON.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bonet, Eulalia. 1991. *Morphology after Syntax: Pronominal Clitics in Romance Languages*. PhD dissertation. MIT, Cambridge, MA.

Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Bošković, Željko. 1997. On certain violations of the Superiority Condition, AGRo, and economy of derivation. In *Journal of Linguistics* 33 (2). 227–254.

Bošković, Željko. 2011. Rescue by PF-deletion, traces as (non)-interveners, and the *that*-trace effect. *Linguistic Inquiry* 42. 1–44.

Brown, Roger & Albert Gillman. 1960. The pronouns of power and solidarity. In Thomas A. Sebeok (ed.), *Style in Language*, 235–276. Cambridge, MA: MIT Press.

Cardinaletti, Anna, & Michal Starke. 1999. The typology of structural deficiency. In Henk van Riemsdijk (ed.), *Clitics in Indo-European Languages*. Eurotyp Project Working Group 8. 145–253. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and Parameters in Comparative Grammar*, 417–454. Cambridge, MA: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.

D'Alessandro, Roberta. 2015. Null subject. In Antonio Fábregas, Jaume Mateu & Michael Putnam (eds.), *Contemporary Linguistic Parameters*, 201–226. London: Bloomsbury Academic Publishing.

Dotlačil, Jasee Chomsky 1981); and Hornstein 1999), b. 2007. Why clitics cannot climb out of CP: A discourse approach. In Richard Compton, Magdalena Golędzinowska & Ulyana Savchenko (eds.), *Formal Approaches to Slavic Languages 15*, 76–93. Ann Arbor: University of Michigan Press.

Emonds, Joseph E. 1987. The Invisible Category Principle. In *Linguistic Inquiry* 18. 613–632.

Emonds, Joseph E. 2012. Blackjack! 21 Arguments that agreeing adjectives are derived nominals. In Esther Torrego (ed.), *Of Grammar, Words and Verses: In Honor of Carlos Piera*, 171–200. Amsterdam: John Benjamins.

Emonds, Joseph & Ludmila Veselovská. 2015. The cross-linguistic homes of Mood and Tense. In Henk van Riemsdijk & Marc van Oostendorp (eds.), *Representing Structure in Phonology and Syntax,* 277–314. Berlin: Mouton de Gruyter.

Fanselow, Gisbert & Lenertová, Denisa. 2001. Left peripheral focus: Mismatches between syntax and information structure. *Natural Language and Linguistic Theory* 29 (1). 169–209.

Filip, Hana & Greg Carlson. 1997. Sui generis genericity. *University of Pennsylvania Working Papers in Linguistics* 4 (2). 91–110.

Franks, Steven, & Tracy H. King. 2000. *A Handbook of Slavic Clitics*. Oxford: Oxford University Press.

Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.

Geist, Ludmila. 1997. Russisch *byt*: zwei Kopulae, zwei Kasus, ein integrierter Ansatz. In Peter Kosta & Jens Frasek (eds.), *Current Approaches to Formal Slavic Linguistics*, 169–178. Frankfurt am Main: Peter Lang.

Gilligan, Gary Martin. 1987. *A cross-linguistic approach to the pro-drop parameter*. PhD dissertation. University of Southern California, Los Angeles, CA.

Grepl, Miroslav & Petr Karlík. 1998. *Skladba češtiny*. [Czech syntax]. Praha: Votobia.

Haspelmath, Martin & Matthew S. Dryer. 2005. *World Atlas of Language Structures*. Oxford: Oxford University Press.

Haspelmath, Martin. 2001. The European linguistic area: Standard average European. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals* (vol. 2), 1492–1510. Berlin: Mouton De Gruyter.

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Hudousková, Andrea. 2009. Clitic *se* – Benefit of being late. In Mojmír Dočekal & Markéta Ziková (eds.), *Czech in Formal Grammar*, 89–101. Berlin: Peter Lang.

Jaeggli, Osvaldo & Ken Safir (eds.). 1989. *The Null Subject Parameter*. Dordrecht: Kluwer.

Kayne, Richard S. 1989. Null subjects and clitic climbing. In Osvaldo Jaeggli & Ken Safir (eds.), *The Null Subject Parameter*, 239–261. Kluwer Academic Publishers.

Kayne, Richard S. 1980. Extensions of binding and case marking. *Linguistic Inquiry* 11. 74–96.

Kučerová, Ivona. 2000. *K subjektpredikátové shodě v češtině* [To subject-predicate agreement in Czech]. MA thesis. Charles University, Prague.

Kučerová, Ivona. 2007. *The Syntax of Givenness*. PhD dissertation. MIT, Cambridge, MA.

Kučerová, Ivona. 2012. Grammatical marking of givenness. In Ivona Kučerová & Ad Neeleman (eds.), *Information Structure: Contrasts and Positions*, 72–86. Cambridge: Cambridge University Press.

Kučerová, Ivona. 2014. The syntax of null subjects. *Syntax* 17 (2). 132–167.

Kučerová, Ivona. 2017. On labelling of DP coordinations and the lack of φ-feature resolution in syntactic Agree. In Andrew Lamont & Katerina Tetzloff (eds.), *NELS 47: Proceedings of the 47 Annual Meeting of the North East Linguistics Society*, 211–220. Amherst, MA: GLSA Publications.

Levin, Beth & Malka Rappaport. 1986. The formation of adjectival passives. *Linguistic Inquiry* 17. 623–661.

Longobardi, Giuseppe. 2008. Reference to individuals, person, and the variety of mapping parameters. In Andreas Klinge & Høeg Müller (eds.), *Essays on Nominal Determination: From Morphology to Discourse Management*, 189–211. Amsterdam: Johns Benjamins.

Meyer, Roland. 2004. Superiority effects in Russian, Polish and Czech: Judgments and grammar. In *Cahiers Linguistiques d'Ottawa* 32. 44–65.

Nevins, Andrew & Jeffrey K. Parrott. 2010. Variable rules meet impoverishment theory: Patterns of agreement levelling in English varieties. *Lingua* 120 (5). 113–159.

Noyer, Rolf. 1998. Impoverishment theory and morpho-syntactic markedness. In Steven Lapointe, Diane Brentari & Patrick Farrell (eds.), *Morphology and Its Relation to Phonology and Syntax*, 264–285. Stanford, CA: CSLI.

Panevová, Jarmila & Vladimír Petkevič. 1997. Agreement in Czech and its formal account. In Uwe Junghanns & Gerhild Zybatow (eds.), *Formale Slawistik*, 321–333. Frankfurt am Main: Vervuert Verlag.

Parrott, Jeffrey K. 2015. Gender impoverishmen in Czech, Slavic, and beyond. In Markéta Ziková, Pavel Caha & Mojmir Dočekal (eds.), *Slavic Languages in the Perspective of Formal Grammar*, 215–232. Berlin: Peter Lang.

Perlmutter, David. 1971. *Deep and Surface Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David. 1982. *Paths and Categories*. Cambridge, MA: MIT PhD dissertation.

Pesetsky, David. 2013. *Russian Case Morphology and the Syntactic Categories*. Cambridge, MA: MIT Press.

Rezac, Milan. 2005. The syntax of clitic climbing in Czech. In Lorie Heggie & Francisco Ordonez (eds.), *Clitic and Affix Combinations*, 103–140. Amsterdam: John Benjamins.

Richards, Mark. 2014. Defective Agree, case alterations, and the prominence of person. In Ina Bornkessel-Schlesewsky, Andrei Malchukov & Mark Richards (eds.), *Scales and Hierarchies*, 173–196. Berlin: Mouton de Gruyter.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. In *Linguistic Inquiry* 17. 501–557.

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics, Incorporation and Defective Goals*. Cambridge, MA: MIT Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction. In Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–56. Cambridge: Cambridge University Press.

Saito, Mamoru. 2007. Notes on East Asian argument ellipsis. *Language Research* 43. 203–227.

Schoorlemmer, Maaike. 1995. *Participial Passive and Aspect in Russian*. Utrecht: Onderzoeksinstituut voor Taal en Spraak (OTS dissertation series).

Schuurman, Matthew. 2016. Deriving superiority effects in multiple WH-fronting without determiners. In *Western Papers in Linguistics / Cahiers linguistiques de Western*: Vol. 3, Article 7.

Short, David 1990. Czech and Slovak. In Bernard Comrie (ed.), *The Major Languages of Eastern Europe*, 101–124. Routledge, London.

Skrabalova, Hana. 2007. Number agreement with coordinate nouns. In Peter Kosta (ed.), *Recent Issues in Formal Slavic linguistics*, 319–332. Frankfurt am Main: Peter Lang.

Skrabalova, Hana. 2008. Coordination: Some evidence for DP and NumP in Czech. In Gerhild Zybatow, Luca Sucsich, Uwe Junghanns & Roland Meyer (eds.), *Proceedings of the 5th meeting of Formal Description of Slavic Languages*, 43–61. Frankfurt am Main: Peter Lang.

Šimík, Radek. 2007. The syntax, semantics, and pragmatics of the focus particle to in Czech. *Proceedings of the 7th meeting of Formal Description of Slavic Languages,* 327–340. Frankfurt am Main: Peter Lang.

Taraldsen, Knut T. 1980. *On the Nominative Island Constraint, Vacuous Application and the That-trace Filter*. Bloomington: Indiana University Linguistic Circle.

Toman, Jindřich. 1980. Weak and strong: Notes on be in Czech. In Grunter Brettschneider & Christian Lehmann (eds.), *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, 305–310. Tübingen: Gunter Narr Verlag.

Toman, Jindřich. 1999. On clitic displacement. In Mila Vulkanova & Lars Hellan (eds.), *Topics in South Slavic Syntax and Semantics*, 205–228. Amsterdam: Johns Benjamins.

Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In Kerstin Schwabe & Susanne Winkler (eds.), *The interfaces deriving and interpreting omitted structures*, 321–339. Amsterdam: Johns Benjamins.

Veselovská, Ludmila. 1994. *Phrasal Movement and X-morphology: Word Order Parallels in Czech and English Nominal and Verbal Projections*. PhD dissertation. Palacký University, Olomouc.

Veselovská, Ludmila. 2001. Agreement patterns of Czech group nouns and quantifiers. In Norbert Corver & Henk van Riemsdijk (eds.), *Semi-lexical Categories*, 273–320. Berlin: Mouton de Gruyter.

Veselovská, Ludmila. 2003. A note about nothing. In Peter Kosta, Joanna Błaszczak, Jens Frasek, Ljudmila Geist & Marzena Żygis (eds.), *Proceedings of the 4th meeting of Formal Description of Slavic Languages,* 745–58. Frankfurt am Main: Peter Lang.

Veselovská, Ludmila. 2004. The extended verbal projection in Czech: Three variants of the verb *be*. In Gerhild Zybatow, Luca Sucsich, Uwe Junghanns & Roland Meyer (eds.), *Proceedings of the 5th meeting of Formal Description of Slavic Languages*, 555–569. Frankfurt am Main: Peter Lang.

Veselovská, Ludmila. 2014. The universal DP analysis in articleless languages: A case study in Czech. In Ludmila Veselovská & Markéta Janebová (eds.), *Nominal Structures: All in Complex DPs*. 12–28. Olomouc: Palacký University.

Veselovská, Ludmila. 2018. *Noun Phrases in Czech: Their Structure and Agreements*. Frankfurt am Main: Peter Lang.

Veselovská, Ludmila. 2020. The structure and categorial characteristics of *pro*. *Zeitschrift für Slawistik* 65 (2). 194–221.

Veselovská, Ludmila & Petr Karlík. 2004. Analytic passives in Czech. *Zeitschrift für Slavistik* 49 (2). 163–243.

Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Wasow, Thomas. 1977. Transformations and the lexicon. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 327–360. New York: Academic Press.

Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford, CA: Stanford University Press.

Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 1. 81–114.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Zanuttini, Raffaella & Judy B. Bernstein. 2008. Micro-comparative syntax in English verbal agreement. In Suzi Lima, Kevin Mullin & Brian Smith (eds.), *Proceedings of the 39th Annual Meeting of the North East Linguistic Society*, 839–860. Amherst: GLSA, University of Massachusetts.

Dobrinka Genevska-Hanke

6 Subject realization in Bulgarian, a consistent null subject language: Theoretical issues and empirical facts

1 Introduction

Bulgarian is a South-Slavic language, in which null referential (and definite) pronominal subjects are licensed in main and subordinate finite clauses (Bojadžiev, Kuzarov & Penčev 1999; Popov 1998; Genevska-Hanke 2019). This has been the case since the time of Old Bulgarian (Mirchev 2000). In contrast, objects have to be generally expressed (see Rizzi 1986 on null objects).

Languages allowing subjects to remain silent, i.e. unpronounced, have traditionally been referred to as "*pro*-drop" languages since the 80s (see Chomsky 1981, 1982; Rizzi 1982, 1986; and Jaeggli & Safir 1989). Italian and Spanish are perhaps two of the most researched languages of this kind and have recently been labelled "consistent" NSLs (see Holmberg 2005; Holmberg, Nayudu & Sheehan 2009; Roberts & Holmberg 2010).[1] Rich verbal inflection expressing subject-verb agreement is crucial for the licensing of the null subject in these languages. This distinguishes them from languages like Chinese or Japanese, which belong to a distinct class of NSLs, that of topic-drop or radical NSLs (Huang 1984; Roberts & Holmberg 2010), in which verbal inflection for person-number agreement is absent and any argument of the verb can remain silent. Null subjects in these languages are licensed differently; they must be recovered through discourse.

Each class of languages of the typological classification given by Roberts & Holmberg (2010) is associated with particular features. Apart from null definite pronominal subjects and rich verbal inflection, the following properties have come to be traditionally known as the cluster of properties of the Null Subject Parameter: (i) expletive null subjects; (ii) post-verbal subjects; (iii) subject extraction from complement clauses (see Rizzi 1982; Roberts & Holmberg 2010).

The question of apparent differences or microvariation among consistent NSLs has been recently raised since some languages seem to be more restrictive than others as to the implementation of the descriptive properties given above

---

1 The term "*pro*-drop" goes back to Chomsky (1981) and Rizzi (1982) within the generative framework, in which parameters are the locus of cross-linguistic variation. Some of the languages traditionally classified as *pro*-drop are viewed as partial NSLs according to the more recent classification of Roberts & Holmberg (2010).

(e.g., Öztürk 2001; Prentza & Tsimpli 2013; Di Domenico & Baroncini 2018; Genevska-Hanke 2019). Some consistent NSLs share properties with languages from other classes, in particular with non-NSLs like English and French as well as partial NSLs like Finnish and Russian. This is true for Bulgarian, which, therefore, cannot be neatly fit into the class of consistent NSLs (Genevska-Hanke 2019). This casts some doubt on the typological classification given by Roberts & Holmberg (2010).

The present chapter will contribute to our understanding of the properties of consistent NSLs in general, and of Bulgarian, in particular. It provides theoretical, empirical and typological arguments in support of its status as a consistent NSL, despite the fact that some of its features pose a problem to this classification. However, the presence of these features is insufficient to negatively affect the availability of null subjects in contrast to the features that induced the language change attested in French, Russian, and more recently Brazilian Portuguese (Roberts 1993, 2014; Müller 2006; Kato 2000). Finally, it is argued that some modification of the classification criteria or the descriptors of the distinct language classes is needed, so that all languages can be adequately placed in one class or another.

2 Null subject theory and issues of classification

2.1 Some theoretical assumptions on null subjects

This section gives a brief theoretical overview on null subjects in consistent NSLs, including the origins of Null Subject Theory (Chomsky 1981, 1982; Rizzi 1982, 1986; Huang 1984; Jaeggli & Safir 1989). Numerous works have arisen ever since, an adequate review of which is beyond the scope of this chapter (but see Roberts & Holmberg 2010; Barbosa 2011). Accordingly, the discussion will focus on a selection of works, in particular those posing difficulty to the task of classifying Bulgarian as a consistent NSL.

From a cross-linguistic point of view, pronominal subjects can either be realized as overt pronouns or remain unexpressed as null pronouns (*pro*'s) in the clausal structure. Reference is made to the grammatical subject-predicate relation so that null subjects are viewed as empty pronominal categories, lacking phonological (and sometimes semantic) content present in syntax, in subject position (Rizzi 1986, 2005). However, it should be noted that some researchers view verbal inflection itself as the element carrying the features in question (see, e.g., Borer 1986; Barbosa 1995; Pollock 1997; Alexiadou & Anagnostopoulou 1998; Platzack

2004) and that yet others do not explicitly associate zero elements with structural positions in general (see Franks 1990; Weiss 2013).

In the present chapter, it is maintained that null pronouns are present in the core syntactic structure. If the other option is chosen (and only inflection is involved), null subjects are limited to being discourse topics much in the sense of topic-drop (Frascarelli 2007; Frascarelli & Hinterhölzl 2007). This is problematic for several reasons, one of which being the fact that the notions of subject and topic are collapsed (see, e.g., Rizzi 2005; Cruschina 2008).

In consistent NSLs, overt and null pronominal subjects alternate and clauses with null subjects are grammatical, compare the Bulgarian examples (1a) and (1b).[2]

(1) a. ___ *znam*   *če*   ___ *idva.*
        (I)   know.PRS.1SG that (he) come.PRS.3SG
        'I know that he is coming.'
     b. *Az znam*   *če*   *toj idva.*
        I   know.PRS.1SG that he   come.PRS.3SG
        'I know that he's coming.'
        (Genevska-Hanke 2019: 3)

The two sentences given in (1a)–(1b) are identical in meaning and reference with regard to the two subjects, which are both referential and definite. The availability of null subjects in a consistent NSL is largely dependent on the richness of verbal inflection but, as will be shown, the fact that a language has rich inflection does not guarantee that null subjects will be licensed. Nevertheless, the fact that a language has both rich inflection and null subjects has related grammatical consequences – a strong expectation for the following properties arises: (i) expletive null subjects; (ii) post-verbal subjects; (iii) subject extraction from complement clauses. Similarly, if one of these descriptive features is lost in the course of historical development, there is a strong expectation that this will impact one of the related properties and induce its loss, a point which will be re-addressed.[3] One example is the relation between null subjects and rich verbal inflection (Rizzi 1993; Roberts 1993; Kato 2000). Besides Italian, and Spanish,

---

2 Note that Bulgarian verbs are morphologically unmarked for present tense, and that 3rd person singular present verbs lack any morphological marking (Bojadžiev, Kuzarov & Penčev 1999). Note that the glosses of the Bulgarian examples nevertheless contain this kind of information in order to improve readability. If the source language is not specified, the examples are from Bulgarian.

3 See Gilligan (1987) for four very robust correlations between the properties of the cluster found to hold cross-linguistically (based on a 100-language sample).

European Portuguese, Greek, Turkish, Hungarian as well as Slavic languages have been described as *pro*-drop languages or NSLs (Rizzi 1986, Roberts & Holmberg 2010). It is noteworthy that all of them have rich verbal morphology.

Apart from purely grammatical conditions, the alternation of overt and null subjects in NSLs in general, and in consistent NSLs, in particular, is further dependent on other aspects of language: the two notions of information structure, "topic" and "focus" determine the choice of the subject.[4] In the context of topic continuity, null subjects are the highly preferred option, while in the context of topic shift or focus, pronominal subjects are realized overtly instead (Samek-Lodovici 1996; Sorace 2005; Frascarelli 2007). This gives rise to interpretative differences for the following two sentences in consistent NSLs. Compare the Italian examples in (2a) and (2b). In (2a) the reference of the overt pronoun is obligatorily disjoint from that of the noun in the main clause, while the null pronoun in (2b) is co-referent with the noun in the main clause.[5]

(2) a. *Il   professore$_i$ ha   parlato dopo che   lui$_{*i/j}$ è   arrivato.*
         the   professor   has   spoken   after that   he   is   arrived
         'The professor started speaking after he had arrived.'
     b. *Il   professore$_i$ ha   parlato dopo che   pro$_{i/*j}$ è   arrivato.*
         the   professor   has   spoken   after that   (he)   is   arrived
         'The professor started speaking after he had arrived.'
         (Italian, Roberts & Holmberg 2010: 7)

Importantly, consistent NSLs do not merely differ from non-NSLs in the availability of null subjects but also in the featural makeup of the overt subject, which is associated with one interpretative option only in consistent NSLs. In contrast, overt subjects in non-NSLs are generally ambiguous between the coreferential and disjoint interpretations.

Focus affects syntax in such a way that it forces the subject to be post-verbal in consistent NSLs like Italian and Spanish (Belletti 2001; Zubizarreta 1998); see the example in (3a) from Italian.

---

4 This is different for the class of radical NSLs, in which discourse factors determine the availability of null subjects. Some researchers propose that these factors are grammatically encoded and suggest that they originate in the grammar itself (e.g., Rizzi 1997; É.Kiss 1998; Belletti 2001, 2009; Aboh 2010).

5 The example has been slightly altered and indices have been supplied to improve the mapping of the two distinct interpretations.

(3) a. *Ha   parlato   Gianni.*
          has   spoken   Gianni
          'Gianni spoke.'
          (Italian, Belletti 2001: 163)

These distinct cross-linguistic patterns are formal options, constrained grammatically and pragmatically, linked to properties of focus expression, and they are termed "answering strategies" (Belletti 2007, 2008). Post-verbal subjects are one of the descriptive features of consistent NSLs.[6]

However, the relation between the availability of post-verbal subjects and null subjects in consistent NSLs cannot be as tight as previously assumed (see Belletti 2009). The presence of *pro* in a language is a necessary but insufficient condition for the availability of post-verbal subjects. The language has to license *pro* in pre-verbal position, which enables the post-verbal subject to move to the post-verbal position for interpretation. The post-verbal position is a focus position in the "VP periphery" in cartographic terms (Belletti 2001). Belletti (2008) suggests that the syntactic position for new information focus is parametrized in consistent NSLs so that some of these languages use a pre-verbal position and others a post-verbal one. In Hungarian, for instance, the pre-verbal position for new information focus is used (but note that this strategy is different form the English one in a structural sense, although the word order of the two is the same).[7]

Rich verbal inflection, which licenses *pro*, is the most prominent feature of the cluster. The relation between rich verbal inflection and the availability of null pronominal subjects is known as Taraldsen's generalization (Taraldsen 1978). The verbal paradigm of Spanish present tense consists of the stem and six distinct verbal endings for the six person-number distinctions (given as (22) in Section 3.1). Jaeggli & Safir (1989) propose the principle of "morphological uniformity" to account for a great deal of the cross-linguistic data (see also Huang 1984). According to this principle, null subjects are available in inflectional systems with only derived (as in Italian) or only underived forms (as in Chinese). Importantly, there

---

6 The availability of post-verbal subjects is traditionally known as "subject-verb inversion" or "free inversion". However, the former term is somewhat imprecise and the latter is rather incorrect, since the kind of inversion is not arbitrary but discourse-dependent, as previously mentioned (see Belletti 2009 for details).

7 According to Belletti, focalization in English is of the type "in-situ" since a structural focus position is unavailable in the language. In addition, some non-*pro*-drop languages like French exploit the post-verbal verbal position in combination with an overt expletive in a cleft construction. Details on post-verbal subjects and on the other properties of the cluster for Bulgarian follow in Section 3.

can be no mixture of the two – derived and underived forms (morphologically complex forms and bare stems) in order for a language to be *pro*-drop. Notice though that there are exceptions, e.g. Mainland Scandinavian lacks subject-verb agreement but still null subjects are disallowed. A similar idea is captured by the "Pro Generalization" proposed by Müller (2006). What follows is his formulation of "impoverishment rules" for system-wide syncretisms in the sense of systematic homonymy of inflectional markers. Note that only one impoverishment rule is claimed enough to block *pro* licensing, for instance the two identical forms for 2nd and 3rd person singular present tense in Icelandic (see (22) in Section 3.1). Importantly, rich inflection does not guarantee *pro*-drop and vice versa (Rizzi 1986).

The licensing of expletive null subjects is another property that has been associated with consistent NSLs. In particular, expletive subjects are typically null in these languages, as the Bulgarian example in (4) indicates.

(4) __ *Vali.*
      (it) rain.PRS.3.SG
      'It is raining.'
      (Bojadžiev, Kuzarov & Penčev 1999: 596)

Expletive subjects are semantically empty and often non-referential. These subjects have to remain unexpressed in consistent NSLs like Italian, Catalan and Spanish, but note that overt expletives are marginally licensed in some languages (see Franks 1990; Rizzi 1993). This is the case for Galician Portuguese, Serbo-Croatian, Polish and Bulgarian (Raposo & Uriagereka 1990; Franks 1990; Bojadžiev, Kuzarov & Penčev 1999). Franks (1990) suggests that for South and West Slavic this can be attributed to the specific nature of the expletive constructions in these languages.

The last descriptive property typical of consistent NSLs is extracting the subject of embedded declaratives or questions without yielding ungrammaticality, as in (5).[8] This property is also referred to as "absence of the *that*-trace filter" (see Perlmutter 1971).

(5) *Koj kaza, če e napisal tazi kniga?*
      who say.PST.3.SG that be.PRF written this book
      'Who did you say wrote this book?'
      (Genevska-Hanke 2019: 76)

---

8 The case of the embedded declarative has been referred to as "resumptive strategies" in the literature. Some researchers have viewed the possibility of the null subject *pro* to act as a resumptive pronoun as a distinct feature of *pro*-drop languages (Jaeggli & Safir 1989).

The descriptive status of this property in consistent NSLs has been questioned because of some cross-linguistic evidence suggesting that these effects are absent even in some non-NSLs, such as some dialects of Mainland Scandinavian, English and Arabic (Sobin 1987; Lohndal 2009; Kenstowicz 1989), and in Norwegian (Rizzi 1993) because the structure of finite clauses does not show these effects (due to the VSO word order pattern) in some NSLs like Welsh and Irish (Roberts & Holmberg 2010). On the other hand, in some non-*pro*-drop languages like Norwegian, that-trace violations are tolerated (Rizzi 1993).

Altogether, it becomes apparent that the presence of these descriptive features in a language does not guarantee that it is a consistent NSL, pointing to problems with classification. This has led to diverging opinions on the nature of certain languages and various related proposals.[9] More recently, additional features for distinguishing types of NSLs have been suggested (e.g., Holmberg 2005, 2010; Roberts & Holmberg 2010; Holmberg & Roberts 2013).

2.2 Additional features of relevance and problems of classification

Microvariation within one language class induces featural overlap with other NSL classes as to the availability of null subjects. This overlap of features across language classes concerns not only the descriptors of consistent NSLs that were mentioned above but also additional features. One of the additional features that have been assumed to be indicative of consistent NSLs is that subject-doubling constructions are hardly ever attested in these languages (Barbosa, Duarte & Kato 2005).

Another concerns recent evidence for the following: while Italian and Greek are stricter in the choice of null subjects in topic continuity (recall (2b)), other consistent NSLs like Spanish, Turkish and Bulgarian allow overt subjects in the same kind of contexts more often (Sorace 2005; Di Domenico & Baroncini 2018; Prentza & Tsimpli 2013; Genevska-Hanke 2019).[10] This is suggestive of variation

---

9 For instance, Russian and Brazilian Portuguese have been classified in more than one way as to the language class they belong to. Russian is described as a *pro*-drop language by Demjjanov & Strigin (2000), Perlmutter & Moore (2002) as well as Müller (2006). However, Franks (1995), Avrutin & Rohrbacher (1997) as well as Junghanns (2005) view its null subjects as instances of topic-drop, like in Chinese. In a similar vein, Roberts & Holmberg (2010) classify Brazilian Portuguese as a partial NSL, while Kato (2000) and Belletti (2008) refer to it as a non-NSL. See also Öztürk (2001) for classifying Turkish as a radical NSL.

10 The assumption that a null pronoun should be chosen over an overt pronoun is grounded in what has come to be known as the "Avoid Pronoun Principle" in the generative literature (Chomsky 1981: 65).

in the frequency of overt vs. null subjects within the class of consistent NSLs. The overlap with the class of partial NSLs becomes apparent also in Holmberg's claim that "null subjects are used more in consistent NSLs than in partial NSLs" (Holmberg 2010: 91). For this reason, the properties of Bulgarian will be checked against the additional features that have been proposed for differentiation between the two language classes of consistent and partial NSLs.

Holmberg (2005), Roberts & Holmberg (2010) as well as Holmberg (2010) propose further criteria to distinguish between consistent and partial NSLs.[11] One concerns definite pronominal null subjects (Holmberg 2010) – in partial NSLs, no 3rd person null subjects are licensed in main clauses at all, while 3rd person null subjects in subordinate clauses need to be licensed by an antecedent in the corresponding main clause. 1st and 2nd person subjects can be null in finite matrix clauses. This is seen as a major difference to the state of affairs in consistent NSLs, where null subjects of this kind are unrestricted in person. However, there seems to be some counter-evidence to that, which blurs the distinction between the two classes of languages. Compare the Russian example (6) and the Finnish example (7).

(6) Net,   ___   *ne kupil*   ___.
    No   (he)   not bought   (it)
    'No, he did not buy it.'
    (Russian, Franks 1990: 91)

(7) ___   *leikkasi*   *peukaloa.*
    (s/he)   cut.PST.3SG   finger.ACC
    'S/he was cutting her/his finger.'
    (Finnish, Dal Pozzo 2011: 49)

A second difference concerns impersonal pronominal null subjects, for which "the situation is, in a sense, reversed" (Holmberg 2010: 92). In partial NSLs, "generic pronouns can, and must, be null" (Holmberg 2005: 539), while "what we do not find in any consistent NSL is a null 3SG indefinite subject" (Holmberg 2010: 92). Consistent NSLs are taken to "always resort to other strategies than a null 3SG pronoun to express the generic inclusive reading for the subject of a finite clause", an overt pronoun, a null 2SG pronoun, a special clitic *si/se* or verb form (medio-) passive (Holmberg 2010: 93).[12] A related assumption is that if the subject

---

11 See also Holmberg & Roberts (2013).
12 Here a difference is made between "generic" and "arbitrary" so that the former means 'one'/'you', referring to people in general including the speaker, while the latter corresponds to "they" referring to people in general but excluding the speaker.

of a generic sentence in a consistent NSL is realized overtly, it must be definite (Holmberg 2010).

A further property that distinguishes consistent NSLs from partial NSLs is related to contexts in which null subjects might be optional in partial NSLs but obligatory in consistent NSLs (contexts where the referent of the main clause and the subordinate clause are identical), like (8), and contexts in which null subjects might not be licensed in partial NSLs but are licensed in consistent NSLs (the referent of the main clause must be disjoint from the referent of the subordinate clause as in (9)).[13, 14]

(8) a. *Ram$_i$ mhanala ki* (*tyani$_i$*) *ghar* *ghetla*.
Ram said that he house bought
'Ram said that he bought a house.'
(Marathi, Holmberg 2010: 91)

  b. *Gianni$_i$ dice che* (**lui$_i$*) *vuole comprare una macchina*.
Gianni says that he wants buy a car
'Gianni says that he wants to buy a car.'
(Italian, Holmberg 2010: 91)

(9) a. *Juha$_1$ ei ole sanonut mitään, mutta Pauli$_2$ sanoo*
Juha NEG.3SG be say.PTCP thing.PART but Pauli say.3SG
että *pro$_1$ haluaa ostaa uuden auton.
that (he) want.3SG buy.INF new car
'Juha$_1$ hasn't said anything, but Pauli$_2$ says that he$_1$ wants to buy a new car.' (Finnish, Holmberg 2010: 92)

  b. *Gianni$_1$ non ha detto niente, ma Paolo$_2$ ha detto*
Gianna not PRF say.PTCP nothing but Paolo PRF say.PTCP
che pro$_1$ *vuole comprare una macchina nouva*.
that (he) want.3SG Buy a car new
'Gianni$_1$ hasn't said anything but Paolo$_2$ says that he$_{1/2}$ wants to buy a new car.' (Italian, Holmberg 2010: 92)

Yet another difference suggested to distinguish between the two language classes concerns pronouns denoting inanimate subjects. These are defined as null in consistent NSLs but overt or null if controlled in partial NSLs (see Holmberg &

---

13 Recall also the Italian example (2b), for which the null subject in the subordinate clause was the highly preferred option.
14 In his examples, Holmberg (2010) represents the empty subject with the empty set symbol (Ø). For the sake of uniformity, Ø has been replaced here with pro.

Sheehan 2010). However, as the authors acknowledge, such pronouns might be sometimes overt in consistent NSLs, as for instance in Arabic.

The discussion so far has shown that there are some problems of classification and that things are not as clear-cut as suggested in the literature. In the following section, language change with regard to *pro*-drop will be briefly addressed since the way languages develop can be informative for classification matters.

2.3 Patterns of historical development in relation to *pro*-drop

Looking at language change and *pro*-drop, it becomes evident that there exists more than one way for a language to switch from *pro*-drop to non-*pro*-drop. The loss of inflectional richness in Brazilian Portuguese has affected the null subject property in such a way that the rates of null subjects have dropped from 74% to 20% within a century (Barbosa, Duarte, & Kato 2005).[15] In Russian, impoverishment of the inflectional paradigm (a derived one, as described by Müller, concerning the loss of the copula and in a second step that of the formally-identical auxiliary in periphrastic tenses) has led to a decrease in the use of null subjects (Müller 2006).[16] Looking at the historical development in French, a different picture arises: two processes took place – a shift in the pronominal paradigms related to the strength of the pronouns in the sense of Cardinaletti & Starke (1994) and a simplification of the agreement paradigm, followed by a loss of the grammatical contexts for null subjects (Roberts 2014). Crucially, null subjects existed for three more centuries after the impoverishment of verbal inflection. Therefore, Roberts suggests that two ways of losing null subjects should be differentiated since the cause for the language change was distinct for each of these two languages. He further proposes three stages with regard to the developmental path in question:

Stage I: Italian, European Portuguese, early French
Stage II: Brazilian Portuguese, intermediate-period French
Stage III: future Brazilian Portuguese, Modern French

What becomes apparent is that the different developmental stages seem to correspond to distinct language classes. Roberts further states that Middle French is comparable to Brazilian Portuguese. This is suggestive of the fact that Brazilian Portuguese and Modern Russian might not have yet reached their final developmental stage, that of

---

15 For a similar pattern, see van Gelderen (2013) for Old English.
16 A rate of 40% of null subjects for contemporary Russian is reported in Genevska-Hanke (2019: 116).

non-NSLs. This observation is relevant for classification matters and raises the question of whether its current classification is justified (recall that Brazilian Portuguese is classified differently in the literature, as a partial-NSL or as a non-NSL). In a similar vein, Dal Pozzo (2011) suggests that Finnish might be developing into a non-NSL due to current changes related to its pronominal paradigms, in particular to their strength. This taken together with the difficulties in disentangling the two language classes raises the question whether partial NSLs should be a class of their own.

As previously indicated, when one descriptive feature of *pro*-drop languages is lost in the course of language change, there is a strong expectation that this will impact (some of) the other *pro*-drop properties as well. For Brazilian Portuguese, this has been attested for the post-verbal subject. Instead of being the highly preferred answering strategy in new information focus contexts as in Italian, the post-verbal subject is currently implemented in the minority of these contexts as one of four strategies in a fairly mixed pattern (Guesser 2007).

3 The null subject property of Bulgarian

3.1 Bulgarian subjects and the properties relevant for classification

3.1.1 Null subjects

Bojadžiev, Kuzarov & Penčev (1999) describe Bulgarian as a *pro*-drop language on the basis of the properties of the traditional cluster. Genevska-Hanke (2019) classifies it as a consistent NSL according to the classification of Roberts & Holmberg (2010). Its null referential subjects are not restricted as to person/number distinctions, tense and clausal type, as exemplified by (10) and (11), previously given as (1a) and (5), and recall (12) for a null quasi-argumental expletive, previously given as (4).

(10)   ___   *Znam*   *če*   ___   *idva.*
     (I)   know.PRS.1SG   that (he) come.PRS.3SG
     'I know that he is coming.'
     (Genevska-Hanke 2019: 3)

(11) *Koj*   *kaza,*   *če*   *e*   *napisal tazi kniga?*
    who   said.PST.3.SG that PRF written this book
    'Who did you say wrote this book?'
    (Genevska-Hanke 2019: 76)

(12) __   *Vali.*
      (it) rain.PRS.3.SG
      'It is raining.'
      (Bojadžiev, Kuzarov & Penčev 1999: 596)

3rd person null subjects are available in main clauses. In subordinate clauses, 3rd person null subjects do not have to co-refer with an antecedent in the matrix clause (Bojadžiev, Kuzarov & Penčev 1999), see (13). This is in line with what is assumed for consistent NSLs and different from what is assumed for partial NSLs (Roberts & Holmberg 2010).

(13) *Petar$_i$ kaza   če   pro$_{i/j}$ e   spal.*
      Peter   say.PST.3SG that (he)   PRF sleep.PTCP
      'Peter$_i$ said that he$_{i/j}$ had slept.'
      (Bojadžiev, Kuzarov & Penčev 1999: 598)

For generic contexts, the following three strategies are used in Bulgarian: the impersonal pronoun *se* in combination with a null subject (arbitrary *pro*), according to Bojadžiev, Kuzarov & Penčev (1999: 609), see (14); a null pronoun with indefinite reference, (15a), or an overt pronoun with indefinite reference, see (15b). These strategies are used interchangeably but note that *se* is blocked in clauses with reflexive verbs so that in some cases the indefinite interpretation cannot be expressed using *se*.[17]

(14) pro$_{arb}$ *da se   puši   e   zabraneno.*
            to   RFL smoke be.PRS.3SG prohibited
      'Smoking is prohibited.'
      (Bojadžiev, Kuzarov & Penčev 1999: 609)

(15) a. ___ *Govoriš   mu,   ___ ne te   sluša.*
          (2SG) talk.PRS.2SG him.CL (3SG) not you listen.PRS.3SG
          'One/you talk(s) to one/him, but nobody/he (doesn't) listen(s).'
      b. *Ti   mu   govoriš   toj ne te   sluša.*
          you him.CL talk.PRS.2SG he   not you listen.PRS.3SG
          'One/you talk(s) to one/him, but nobody/he (doesn't) listen(s).'
          (Bojadžiev, Kuzarov & Penčev 1999: 600)

---

17 The examples are kept in their original form as far as the notation of the null subject is concerned. See Rizzi (1986) for a similar Italian example. In addition, note that Fassi Fehri (2009) suggests some modification as to the availability of null generic *pro* in NSLs.

Importantly, the overt pronoun in (15b) is ambiguous between the definite and the indefinite interpretations (Bojadžiev, Kuzarov & Penčev 1999), which contradicts the expectation that the subject of a generic sentence in a consistent NSL cannot be realized overtly. Accordingly, the association of overt pronominal subjects and definiteness might be less tight than assumed. In a similar vein, Barbosa (2011) suggests that Holmberg's claim for an obligatorily referential interpretation for null subjects in *pro*-drop languages is possibly too strong since many *pro*-drop languages do have 3rd person plural null subjects with non-referential interpretation.

3.1.2 Overt subjects

As previously mentioned, overt subjects do not exclusively occur in topic shift and focal contexts but also in contexts of topic continuity as in (16).

(16) *Petar$_1$ kaza če toj$_{1/2}$ e spal.*
Peter say.PST.3SG that he PRF sleep.PTCP
'Peter said that he had slept.'
(Bojadžiev, Kuzarov & Penčev 1999: 598)

This example presents counter-evidence to the assumption that an overt pronominal subject in the subordinate clause cannot co-refer with the subject of the matrix clause in consistent NSLs. It is particularly telling with regard to the contexts in which null subjects are optional in partial NSLs but obligatory in consistent NSLs, as in (8), since the overt subject of the subordinate clause can co-refer with the subject of the main clause. Note that the use of overt subjects in contexts of topic continuity is attested in Old Bulgarian (Haralampiev 2001). Crucially, this is different from the state of affairs in Italian and Greek, in which the null pronoun is the highly preferred option. However, the use of overt subjects in topic continuity has been attested in other consistent NSLs like Spanish (Prentza & Tsimpli 2013). The cross-linguistic difference in question has been referred to as an instance of microvariation across consistent NSLs and it has been attributed to the different scope and strength of overt subjects across consistent NSLs (see Genevska-Hanke 2019).

According to Cardinaletti & Starke's (1994) typology of structural deficiency for pronominal elements, weak overt subjects and *pro* share the same features, which allows for their alternation according to their proposal. Weak pronouns and clitics are deficient elements in comparison to strong pronouns and the distinction between these three pronominal classes triggers asymmetries in the domains

of syntax, morphology, semantics and phonology. These asymmetries are attributed to the lack of certain features of the functional projection of the pronominal element. One example is that only pronouns referring to human entities can be used in coordination structures (strong pronouns). Weak pronouns are generally more restricted in their syntactic distribution than strong pronouns. From a semantic point of view, pronouns referring to non-human entities are weak. Phonologically, only strong pronouns carry word stress. Languages differ in the type and number of pronominal paradigms for subjects and objects. For instance, Italian has two formally different pronominal paradigms for subjects. Note that according to Cardinaletti (2004) the pronominal paradigms in Italian are undergoing a change, and the pronouns *lui* and *lei* are turning weak, taking up the slots of the archaic *egli* and *ella* at the same time. This goes together with an increase in the use of overt subjects, see also Frascarelli (2007) for related evidence. In contrast, the English pronoun *he* can be both weak or strong and so does the corresponding Bulgarian pronoun *toj* (Genevska-Hanke 2019). Genevska-Hanke suggests a pronominal paradigm for Bulgarian that is weak and formally identical to the one for strong pronouns. This assumption is supported by the fact that overt pronouns are commonly used to refer to inanimate entities, another feature that existed already in Old Bulgarian (Haralampiev 2006), see (17).

(17) *Botušăt   j   se   săbuva   sam –* pro *mnogo e*
     boot.the   her   RFL   take.off.PRS.3SG self (it) much   be.PRS.3SG
     širok. toj   trudno   se   săbuva   kogato   e   mokăr.
     wide   he   hard   RFL   take off.PRS.3SG when   be PRS.3SG wet
     'Her boot slips off itself – it's too wide. It's hard to take off, when it's wet.'
     (Bojadžiev, Kuzarov & Penčev 1999: 603)

The same observation holds for Arabic dialects, described as consistent NSLs (Kenstowicz 1989; Fassi Fehri 2009) despite the fact that the use of overt pronouns with inanimate reference is viewed as a property of partial NSLs.

With regard to subject doubling, (18a) is an instance without dislocation, while (18b) shows the dislocation of the lexical subject (Bojadžiev, Kuzarov & Penčev 1999). Note that subject doubling is another feature inherited from Old Bulgarian (Haralampiev 2001). Importantly, it is a feature which is claimed to be hardly ever attested in consistent NSLs by Barbosa, Duarte & Kato (2005).

(18) a. *Studentăt   toj šče   dojde   utre.*
        student.the he   will come   tomorrow
        'The student will come tomorrow.'
        (Bojadžiev, Kuzarov & Penčev 1999: 565)

b. *Toj šče dojde utre studentăt.*
he will come tomorrow student.the
'The student will come tomorrow.'
(Bojadžiev, Kuzarov & Penčev 1999: 565)

As previously mentioned, post-verbal subjects are possible in Bulgarian, see (19).

(19) *Idva Ivan.*
come.PRS.3SG Ivan
'Ivan is coming.'
(Genevska-Hanke 2019: 78)

However, if the extent of the use of post-verbal subjects in the context of new information focus is taken into account, in contrast to Italian, Bulgarian only makes use of the post-verbal position in 30% of all cases, according to the data obtained with a test on elicited production of new information focus subjects (Genevska-Hanke 2019, the details follow in Section 3.2.2). The main answering strategy or pattern in these contexts is a pattern with a pre-verbal subject (Genevska-Hanke 2019).

Except for the licensing of null expletives in the majority of cases, Bulgarian allows for some expletives to be overt, as shown below.

(20) *To naistina vali.*
it really rain.PRS.3SG
'It is really raining.'
(Bojadžiev, Kuzarov & Penčev 1999: 602)

Note that the overt element in the slot of a quasi-argumental expletive in (20), which is formally identical to the expletive and referential pronoun *to*, is described as "probably a particle in this case" (Bojadžiev, Kuzarov & Penčev 1999). However, in other cases it is more problematic to classify the element filling the subject slot as a particle; see a non-argumental expletive in (21).

(21) *To me e sram.*
it me be.PRS.3SG shame
'I'm ashamed.'
(Genevska-Hanke 2000: 32)

As indicated elsewhere, the marginal use of overt expletives seems to be common for Slavic languages (Franks 1990). For Bulgarian, this use might be attributed

to the fact that Bulgarian *to* possibly has some kind of demonstrative remnants (Hentschel, p.c.), as has been suggested for English *they* and *them* (Cardinaletti & Starke 1994). Overt expletives are also attested in Galician Portuguese (Rizzi 1993). On the other hand, null expletives are used in a language like Russian, which has a more restricted distribution of null subjects in comparison to other Slavic languages like Polish and Bulgarian. Thus, it can be concluded that the availability of null or overt expletives cannot be used as a distinguishing feature of NSL subclasses to a sufficient extent.

3.1.3 A note on the Bulgarian verbal paradigm

Turning to the syncretism in the verbal paradigm for past tense in Bulgarian, it can be observed that the forms of 2nd and 3rd person singular fall together in the aorist and the imperfect tenses, see Bulgarian B in (22). This uniformity is system-wide as it cannot be tracked back to the underspecification of individual inflectional markers or accidental homonymies. Accordingly, Bulgarian would not be classified as *pro*-drop if Müller's (2006) approach were applied, because one system-wide syncretism should be enough to induce *pro*-blocking for the system, as previously noted and as assumed for the non-NSL Icelandic (which is classified correctly). Note that this syncretism has already existed in Old Bulgarian (Mirchev 2000). In addition, identical endings leading to ambiguities are also attested in Spanish (Jaeggli & Safir 1989). Accordingly, some marginal morphological overlap seems tolerable for *pro*-drop language systems, as suggested by Rizzi (2005) as well as Roberts (1993, 2014). According to Roberts, the degree of impoverishment of inflectional richness is more important than its mere existence and zero endings should be counted as distinct forms of the paradigm in addition to inflectional endings as long as enough distinctions are made. Note that the Bulgarian verbal paradigm for present tense of the verb *idvam* in the meaning of 'to come', given in (22) (Bulgarian A), contains one zero ending – the one for 3rd person singular.[18] In other words, the verb remains unmarked for person (Bojadžiev, Kuzarov & Penčev 1999). This is relevant for theoretical proposals and classification matters.

---

18 Inflectional endings for the present paradigms on the left are given in bold, identical forms for the past paradigms on the right are underlined.

(22)

Spanish[19]		Bulgarian A		Icelandic[20]	
PRS.1SG	habl**o**	PRS.1SG	idva**m**	PRS.1SG	slepp
PRS.2SG	habl**as**	PRS.2SG	idva**sh**	PRS.2SG	sleppur
PRS.3SG	habl**a**	PRS.3SG	idva	PRS.3SG	sleppur
PRS.1PL	habl**amos**	PRS.1PL	idva**me**	PRS.1PL	sleppum
PRS.2PL	habl**ais**	PRS.2PL	idva**te**	PRS.2PL	sleppið
PRS.3PL	habl**an**	PRS.3PL	idva**t**	PRS.3PL	sleppa

Bulgarian B	
PST.1SG	hlaznah se
PST.2SG	hlazna se
PST.3SG	hlazna se
PST.1PL	hlaznahme se
PST.2PL	hlaznahte se
PST.3PL	hlaznaha se

Summing up so far, the following features seem problematic for classifying Bulgarian as a consistent NSL: (i) the peculiarities of its verbal paradigm (syncretism and zero ending), (ii) the possibility of having overt expletives, (iii) the use of pre-verbal subjects in new information focus contexts, (iv) subject doubling, (v) overt subjects in contexts of topic continuity, (vi) contexts in which null subjects are optional in partial NSLs but obligatory in consistent NSLs (since they are also optional in Bulgarian), and (vii) the use of overt subjects with inanimate reference. With regard to the features inherited from Old Bulgarian, the syncretism of the Bulgarian past tense paradigm, subject doubling, overt subjects in topic continuity and the use of overt subjects with inanimate referents, it is implausible that their presence in the language has any impact on the null subject property, because they are stable (given that null subjects have co-existed for over a millennium). In addition, some of these features have been attested in other consistent NSLs. This calls for the reconsideration of the criteria used to identify this language class.

3.2 Empirical support for Bulgarian being a consistent NSL

3.2.1 Null subject rates

Given the problematic nature of classifying Bulgarian as a consistent NSL on the basis of (traditional) classificational criteria, it would be highly informative to look

---

19 The Spanish paradigm is from Jaeggli & Safir (1989: 27).
20 The example is from Müller (2006: 11).

at the extent to which null subjects are used by monolingual speakers. Note that this kind of data is crucially needed since it goes beyond the descriptive character of observations, providing details on the extent to which given features are actually implemented in a language. Genevska-Hanke (2019) analyzed the rates of null and overt subjects in a corpus of contemporary Bulgarian (2000 utterances). The corpus consists of spontaneous speech production of monolingual adults from Eastern Bulgaria. The results reveal that Bulgarian is a consistent NSL with overall high rates of null subjects, 61–73% depending on whether all or only pronominal subjects are counted. In comparison, 67–74% of null subjects are attested for Italian (Lorusso, Caprin & Guasti 2005; Di Domenico & Baroncini 2018), 69% for Greek (Di Domenico & Baroncini 2018) and 78% for European Portuguese (Barbosa, Duarte & Kato 2005).[21] The rates of null pronominal subjects in Bulgarian main clauses reach 70%, while that in subordinate clauses is even higher – 86%, which is a solid indication for the null subject property of consistent NSLs.[22]

In addition, overt subjects in topic continuity contexts were attested in the corpus, but most of these were implemented for a reason – in contexts of copula elision, (23), and the clitic constraint, (24), as described by Genevska-Hanke (2019). In (23), the overt subject is necessary for indicating the relation between the subject and the predicate in the absence of the copula. Without an overt subject clause-initially, the sentence in (24) would be ungrammatical unless movement is involved, which would be more costly in grammatical terms.

(23) ___₁ *kaza*   ___₂ *šče go izhvărlja,* ___₂ *šče go izpădja*
      (he) said.PST.3SG (I) will him throw.out (I) will him chase
      na ulitsata i tja raztrevožena.
      to street.the and she worried
      'He says: I will throw him out, I will chase him away! And she's worried.'
      (Genevska-Hanke 2019: 94)

---

21 Note that these rates are not fully comparable due to the kind of data and type of sentences considered. For instance, in the Bulgarian data, two overt subjects were counted with each instance of subject doubling. This results in slightly higher rates for overt and slightly lower rates for null subjects.

22 The rates for main and subordinate clauses were counted for the data of one speaker. Note that a higher rate of null subjects is generally indicative of *pro*-drop so that a difference between consistent and partial NSLs is one of degree.

(24) *Te   se   poznavat   ot   mnogo godini.*
      they RFL know.PRS.3PL from   many   years
      'They have known each other for many years.'
      (Genevska-Hanke 2019: 91)

3.2.2 Post-verbal subjects in new information focus contexts

Being one of the descriptive features of consistent NSLs, the implementation of post-verbal subjects in the language can aid classification. Again, the mere possibility of having post-verbal subjects is rather insufficient – what is needed is information on the frequency of their use in combination with the implementation of a parametrized focus position. Post-verbal subjects are used in the vast majority of answers to new information focus questions in Italian (98%) as one of only two strategies (Belletti & Leonini 2004). This is different for Brazilian Portuguese and Finnish, which are classified as partial NSLs (Guesser 2007; Dal Pozzo 2011). The patterns of these two languages are mixed, involving three or four answering strategies each. For Finnish, these are SV answers (83%), VS answers (8%) and clefts (9%).[23] SV is the highly preferred strategy. In the cases where VS was employed, another overt phrase necessarily filled the pre-verbal position. The reported rates for the consistent NSL Italian and the partial NSLs Brazilian Portuguese and Finnish are comparable since the same test was used, the one developed for Italian by Belletti & Leonini (2004). This video test targets the elicitation of spontaneous verbal new information subject clauses (the participants answer questions of the kind "Who called?" to 22 short videos). Genevska-Hanke (2019) applied this test to Bulgarian, extending it to objects in new information focus contexts. The results show that in Bulgarian the low peripheral new information focus position is exploited in 30% of the cases of new information focus subjects.[24] Recall that in consistent NSLs the new information focus position is parametrized so that new information subjects and objects occupy either a pre- or a post-verbal position (Belletti 2008). Bulgarian objects occupy the post-verbal position in 99% of all new information focus contexts. Overall, the Bulgarian pattern is like the Italian one but the extent to which post-verbal subjects are implemented is much lower. DP-internal focalization was used in 67% of the Bulgarian answers, which is the strategy used in English.[25] Passives and clefts were used very marginally so that

---

23 Rates calculated on the basis of data from Dal Pozzo (2011).

24 Note that generational and regional differences were attested.

25 No differences as to type of verb (unaccusative, unergative or transitive) were attested, which is different from the Italian data.

the overall pattern for new information focus subjects has not been considered as truly mixed but rather consisting of two main answering strategies. Crucially, no overt expletives were implemented in the VS answers, which is another solid indication for the null subject property of consistent NSLs. In addition, these results undermine an important difference from partial NSLs, since overt expletives were used in the corresponding clauses in Finnish (Dal Pozzo 2011).

The Bulgarian pattern is thus not only different from the patterns of Brazilian Portuguese and Finnish but also different from that of Italian. However, Italian is the only consistent NSL for which the video test has been applied so far, which allows for the possibility that the uniformity of answers in Italian is rather exceptional, as suggested by Belletti (p.c.). With regard to this, it should be noted that there is data according to which Spanish might be similar to Bulgarian and different from Italian as to the implementation of SV as an answering strategy (Hoot 2012; Vanrell & Fernández 2017).[26]

3.2.3 Preference for null over overt subjects in topic continuity contexts

Genevska-Hanke (2020) reports results on the use of and preference for null subjects in topic continuity contexts in Bulgarian from a corpus of spoken language and an online self-paced reading task. The corpus consists of story-telling and retelling of MAIN material, series of six pictures (Gagarina et al. 2012).[27] The results show that adult monolingual speakers of Bulgarian use overt subjects in contexts of topic continuity in up to 32% of the cases, which is comparable to Spanish (27% reported by Shin & Cairns 2009). Accordingly, null subjects are clearly preferred in these contexts.

The self-paced reading task included items in which either null or overt subjects are used in topic continuity contexts, (25) and (26) exemplify one of each kind.

(25) *Mara   ne   e   stignala daleč, a veče   e   ___   dala*
     Mara   not   PERF got   far   but already PERF (she) given
     vsičkite si   pari   na   vjatara.
     all   her   money   to   wind.the
     'Mara hasn't got far, but has already spent all her money for nothing.'
     (Genevska-Hanke 2020: 40)

---

26 Note that results heavily depend on test design, and spoken vs. written data make a difference.
27 32 participants were tested. Here, reference is made to the adult data only, but note that the original study deals with the comparison of adult and developing grammars.

(26) *Tim e   hodil   čak do Yaponiya, a   toj e   izharčil*
     Tim PERF go.PTCP far to Japan   but he   PERF spend.PTCP
     mnogo malo na patuvaneto.
     too   little on trip.the
     'Tim has gone as far as Japan, spending extremely little for the trip.'
     (Genevska-Hanke 2020: 40)

The reaction times measured after an overt subject compared to those after a null subject imply that the monolingual participants read significantly slower when an overt subject was implemented in topic continuity. This result can be also interpreted as indicative of a preference for null subjects in contexts of topic continuity.

4 General discussion

The main descriptors of consistent NSLs are null referential subjects in finite clauses, unrestricted as to clausal position, rich verbal inflection, null expletives, post-verbal subjects and the absence of *that*-trace effects. For Bulgarian, three of these five features are not implemented the way they are in Italian, the best researched consistent NSL. In relation to rich verbal inflection, the Bulgarian verbal paradigm for past involves one system-wide syncretism and one of the six inflectional endings in the present paradigm is zero. However, one syncretism seems tolerable for *pro*-drop systems and so does a zero ending, especially if six distinctions are present for the six person-number members of the agreement paradigm (see Roberts 1993, 2014). Second, Bulgarian expletives are not always null, but this seems to be the case in other Slavic languages, too (Franks 1990). More generally, the availability of null expletives does not differentiate sufficiently between language classes. Third, Bulgarian uses post-verbal subjects in a way consistent NSLs do in only 30% of all cases – in the sense of exploiting a parametrized position for new information focus. In Italian, this position is exploited in 98% of all cases. However, no other consistent NSLs have been investigated with the test of Belletti & Leonini (2004), so it is possible that more languages are less strict than Italian (Belletti, p.c.). There are results on focus in Spanish that point to the implementation of SV, which might be an indication that the Spanish pattern is similar to the Bulgarian one (Vanrell & Fernández 2017).

Considering the descriptive features of consistent NSLs, the discussion so far points to some obvious limitations as to their classifying potential. A major issue is the fact that for each of these properties some cross-linguistic counter-evidence

exists. Accordingly, the presence of these features in a language may be taken as an indication of the properties of NSLs but does not guarantee it. The cross-linguistic exceptions point to the existence of some degree of tolerance that such null subject systems have (see Rizzi 2005). Importantly, the exceptions go beyond individual languages, one example being the possibility of overt expletive subjects in Slavic languages (Franks 1990). An additional point of discussion was the difference in the extent to which the Avoid Pronoun Principle holds among consistent NSLs, which is subject to microvariation.

With regard to the additional features that have been assumed to distinguish between consistent and partial NSLs, it could be shown that discrimination was rather insufficient due to the cross-linguistic counter-evidence. These features are: (i) the need of an antecedent for 3rd person null subjects in subordinate clauses and their unavailability in main clauses in partial NSLs, (ii) the absence of null generic pronouns in combination with the obligatorily definite reference of overt pronouns in consistent NSLs, (iii) contexts in which null subjects are optional in partial NSLs but obligatory in consistent NSLs (see (8)), (iv) contexts in which null subjects are not licensed in partial NSLs but licensed in consistent NSLs (see (9)), (v) overt pronouns with inanimate reference, and (vi) subject doubling. Recall that subject-doubling is attested both in Brazilian Portuguese and Bulgarian, although these languages are clearly different in terms of their NSL properties. In addition, some consistent NSLs allow overt subjects in topic continuity to a larger extent than others (see property (iii) above).

For many of the traditional descriptors of consistent NSLs and the additional features suggested to distinguish them from partial NSLs, (counter)evidence from Bulgarian has been presented. Based on the descriptive typological data that has been provided and the empirical results from the studies on Bulgarian, it can be concluded that this language is undoubtedly a consistent NSL, despite the presence of the features that are not typical of this language class. The attested overall high rates of null subjects (especially in subordinate clauses), comparable to those in Italian and European Portuguese, the exploitation of a parametrized new information focus position, the null subjects in cleft answers to new information focus questions, and the preference for null subjects in topic continuity contexts in adult and developing grammars are highly indicative of the null subject property in consistent NSLs. As for the use of overt subjects in topic continuity contexts in the corpus of spontaneous speech production, many of these subjects were implemented for grammar-internal reasons. In addition, if a language has a pronominal paradigm for weak overt pronouns, these pronouns freely alternate with *pro* in any grammatical context. Importantly, the presence of overt subjects in topic continuity contexts does not seem to negatively affect the null subject property of Bulgarian, despite the long-term presence of this feature in the language. Accordingly, the

evolvement or existence of weak pronouns is rather a feature more loosely related to the availability of null subjects and might not be Roberts' second way for a consistent NSL to lose its null subjects. Crucially, with regard to the classification, the possibility of having overt subjects in topic continuity contexts does not mean that the language in question is not a consistent NSL or that it is losing its null subject nature, even more so since there are more consistent NSLs that allow for overt subjects in topic continuity contexts. Accordingly, the classification criteria need reconsideration. As for the problems of disentangling partial and consistent NSLs, these might be attributable to the possibility that partial NSLs might not be a genuine language class on their own but an intermediate stage of historical development. With regard to this, recall that Brazilian Portuguese is sometimes classified as a non-null subject language (Belletti 2001). In a similar vein, Dal Pozzo (2011) suggests that Finnish is possibly developing into a non-null subject language.

Thus, it can be concluded that the role of rich agreement is pervasive and that as long as verbal inflection is rich in a consistent NSL, this language will continue licensing null subjects irrespectively of the presence of weak overt subjects in topic continuity contexts and other features that are not typical of consistent NSLs. In contrast, recall that both Brazilian Portuguese and Russian have experienced impoverishment of verbal inflection, which is considered the driving force for the loss of null subjects.

5 Conclusion

This paper provided a detailed description of the features associated with the null subject property of consistent NSLs for Bulgarian and presented empirical data to support the assumption that this language belongs to the class of consistent NSLs, despite the fact that features that are not typical of consistent NSLs are also attested. It could be shown that using the traditional descriptors of consistent NSLs incorrectly excludes not only Bulgarian but also other consistent NSLs from the class of consistent NSLs. Moreover, typological arguments based on cross-linguistic comparisons revealed that there is counterevidence for each of the traditional descriptive properties of these languages (the properties of the cluster). Thus, it is not surprising that some modifications and a looser connection between the descriptors and the language class have been previously suggested (e.g., Roberts 2014 for verbal inflection and Belletti 2009 for post-verbal subjects). If consistent NSLs allow for a marginal overlap with other language classes, this microvariation should be properly accounted for so that the classification can be less rigid overall for a phenomenon of an apparently gradient nature.

With regard to the additional properties that have been suggested to discriminate between partial and consistent NSLs, a similar picture arises – in light of cross-linguistic counter-evidence, the features in question might be present or absent in a language for independent reasons. Thus, some modification of the classification criteria is necessary to ensure the proper inclusion of all consistent NSLs into the class of consistent NSLs. Importantly, ongoing developmental changes and related empirical data should be considered to further justify classification and improve its reliability.

References

Aboh, Enoch O. 2010. Information structuring begins with the numeration. *Iberia: An International Journal of Theoretical Linguistics* 2(1). 12–42.

Alexiadou, Artemis & Elena Anagnastopoulou. 1998. Parametrizing AGR: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539.

Avrutin, Sergej & Bernhard Rohrbacher. 1997. Null subjects in Russian inverted constructions. *Proceedings of the Fourth Annual Workshop on Formal Approaches to Slavic Linguistics*, 32–53. Michigan: Slavic Publications, Ann Arbor.

Barbosa, Pilar. 1995. Null subjects. Cambridge, MA: MIT PhD dissertation.

Barbosa, Pilar. 2011. *Pro*-drop and theories of *pro* in the minimalist program. Part 1: Consistent null subject languages and the pronominal-Agr hypothesis. Part 2: Pronoun deletion analyses of null subjects and partial, discourse and semi *pro*-drop. *Language and Linguistics Compass* 5(8). 551–587.

Barbosa, Pilar, Maria E. Duarte & Mary A. Kato. 2005. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 4. 11–52.

Belletti, Adriana. 2001. Inversion as focalization. In Aafke Hulk & Jean-Yves Pollock (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*, 60–90. Oxford: Oxford University Press.

Belletti, Adriana. 2007. Answering strategies. A view from acquisition. In Sergio Baauw, Frank Drijkoningen & Manuela Pinto (eds.), *Romance Languages and Linguistic Theory,* 19–38. Amsterdam: John Benjamins.

Belletti, Adriana. 2008. New information subjects and the nature of clefts. Ms., University of Siena.

Belletti, Adriana. 2009. *Structures and Strategies*. London: Routledge.

Belletti, Adriana & Ciara Leonini 2004. Subject inversion in L2 Italian. In Susan Foster-Cohen, Michael Sharwood Smith, Antonella Sorace & Mitsuhiko Ota (eds.), *Eurosla Yearbook 4*, 95–118. Amsterdam: John Benjamins.

Borer, Hagit. 1986. I-Subjects. *Linguistic Inquiry* 17. 375–416.

Bojadžiev, Todor, Ivan Kuzarov & Jordan Penčev. 1999. Săvremenen bălgarsky ezik [Modern Bulgarian language]. Sofia: Petar Beron.

Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In Luigi Rizzi (ed.), *The Structure of CP and IP*, 115–165. Oxford: Oxford University Press.

Cardinaletti, Anna & Michael Starke. 1994. The typology of structural deficiency: on the three grammatical classes. *University of Venice. Working Papers in Linguistics* 4(2). 41–109.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris: Dordrecht.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Cruschina, Silvio. 2008. *Discourse-related Features and the Syntax of Peripheral Positions*. PhD dissertation. Cambridge: University of Cambridge.

Dal Pozzo, Lena. 2011. *Testing the acquisition of new information subjects: Evidence from Italian and Finnish*. PhD dissertation. Siena: University of Siena.

Demjjanov, Assinija & Anatoli Strigin. 2000. Case assignment to conceptual structures: the Russian instrumental. In Marcus Kracht & Anatoli Strigin (eds.), *Papers on the Interpretation of Case. Linguistics in Potsdam*, Vol. 10. 75–107. Universität Potsdam, Institut für Linguistik.

Di Domenico, Elisa & Ioli Baroncini. 2018. Age of onset and dominance in the choice of subject anaphoric devices: Comparing natives and near-natives of two null-subject languages. *Bilingual Language Development*. Front. Psychol. doi: 0.3389/fpsyg.2018.02729.

Fassi Fehri, Abdelkader. 2009. Arabic silent pronouns, person and voice. *Brill's Annual of Afroasiatic Languages and Linguistics* 1. 1–38.

Franks, Steven. 1990. On the status of null expletives. *Lingua* 81. 1–24.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential pro. An interface approach to the linking of (null) pronouns. *Natural Language Linguistic Theory* 25. 691–734.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Susanne Winkler & Kerstin Schwabe (eds.), *On Information Structure, Meaning and Form,* 87–116. Amsterdam: John Benjamins.

Gagarina, Natalia, Kunnari, S., Tantele, K., Välimaa, T., Balciuniene, I., Bohnacker Ute, Walters, J. & D. Klop. 2012. MAIN Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics* 56.

Gelderen, van, Elly 2013. Null subjects in Old English. *Linguistic Inquiry* 44 (2). 271–285.

Genevska-Hanke, Dobrinka. 2019. *Overt and Null Subjects in Bulgarian and in L1 Bulgarian-L2 German Interlanguage*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Genevska-Hanke, Dobrinka. 2020. Investigating pronominal knowledge in Late L1 attrition: What comparisons of different types of monolinguals and bilinguals can tell us? Ms., University of Oldenburg.

Gilligan, Gary M. 1987. *A Cross-linguistic Approach to the Pro-drop Parameter*. PhD dissertation, University of Southern California, Los Angeles, CA.

Guesser, Simone. 2007. *Soggetto nullo e focalizzazione del soggetto in Portoghese Brasiliano* [Null subjects and subject focalization in Brazilian Portuguese]. MA thesis. Siena: University of Siena.

Haralampiev, Ivan. 2001. *Istoričeska gramatika na bălgarskija ezik* [A historical grammar of Bulgarian]. Veliko Tarnovo: Faber.

Haralampiev, Ivan. 2006. *Bădeščeto na bălgarskija ezik ot istoričesko gledišče* [The future of Bulgarian from a historical perspective]. Veliko Tarnovo: Faber.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97.

Holmberg, Anders & Michelle Sheehan. 2010. Control into finite clauses in null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–152. Cambridge: Cambridge University Press.

Holmberg, Anders & Roberts, Ian. 2013. The syntax-morphology relation. *Lingua* 130. 111–131.

Hoot, Bradley. 2012. *Presentational focus in heritage and monolingual Spanish*. PhD dissertation. University of Illinois, Chicago, IL.

Huang, James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Jaeggli, Osvaldo & Kenneth Safir. 1989. *The Null Subject Parameter*. Dordrecht: Foris.

Junghanns, Uwe. 2005. Bedingungen für die Argumentrealisierung in Slawischen Schprachen. Projektantrag im Rahmen der DFG-Forschergruppe Grammatik und Verarbeitung verbaler Argumente [Conditions on the realization of arguments in Slavic. Project application within the DFG research group Grammar and Processing of Verbal Arguments]. Leipzig: Universität Leipzig, MPI-CBS, MPI-EV.

Kato, Mary. 2000. The partial *pro*-drop nature and the restricted VS order in Brazilian Portuguese. In Mary Kato & Esmeralda Negrao (eds.), *Brazilian Portuguese and the Null Subject Parameter*, 223–258. Frankfurt: Vervuert.

Kenstowicz, Michael. 1989. The null subject parameter in Modern Arabic dialects. In Jaeggli, Osvaldo & Kenneth Safir (eds.), *The Null Subject Parameter*, 263–275. Dordrecht: Foris.

Kiss, É. Katalin. 1998. Identificational focus versus information focus. *Language* 74. 245–273.

Lohndal, Terje. 2009. *That*-trace in Scandinavian and elsewhere: Variation in the position of C. *Studia Linguistica* 63. 204–232.

Lorusso, Paolo, Claudia Caprin & Maria T. Guasti. 2005. Overt subject distribution in early Italian children. In Alejna Brugos, Manuela R. Clark-Cotton & Seungwan Ha (eds.), *A Supplement to the Proceedings of BUCLD*. Somerville MA: Cascadilla Press.

Mirchev, Kiril. 2000. *Starobǎlgarski ezik* [Old Bulgarian Language]. Veliko Tarnovo: Faber.

Müller, Gereon. 2006. *Pro*-drop and impoverishment. In Patrick Brandt & Eric Fuß (eds.), *Form, Structure, and Grammar. A Festschrift Presented to Günther Grewendorf on the Occasion of his 60th Birthday*, 93–115. Berlin: Akademie Verlag.

Öztürk, Balkiz. 2001. Turkish as a non-*pro*-drop language. In Eser E. Tayland (ed.), *The Verb in Turkish*, 239–260. Amsterdam: Benjamins.

Perlmutter, David. 1971. *Deep and Surface Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Perlmutter, David, and Jimmy Moore. 2002. Language-internal explanation: the distribution of Russian impersonals. *Language* 78: 619–650.

Popov, Konstantin. 1998. *Syntaksis na bǎlgarsija knižoven ezik* [Syntax of the Bulgarian written language]. Veliko Tarnovo: Abagar.

Platzack, Christer. 2004. Agreement and the person phrase hypothesis. *Working Papers in Scandinavian Syntax* 73. 83–11

Pollock, Jean-Y. 1997. *Langage et cognition: introduction au programme minimaliste de la grammaire generative*. [Language and cognition: introduction to the minimalist programme of generative grammar]. Paris: Presse Universitaires de France.

Prentza, Alexandra & Ianthi. M. Tsimpli. 2013. On the optionality in L2 pronominal production and interpretation. What (more) can VP-coordination structures tell us? *Eurosla* 13. 22–46.

Raposo, Eduardo & Juan Uriagereka. 1990. Long-distance case assignment. *Linguistic Inquiry* 21 (4). 505–537.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro. Linguistic Inquiry* 17. 501–57.

Rizzi, Luigi. 1993. A parametric approach to comparative syntax: properties of the pronominal system. *English Linguistics* 10. 1–27.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Heageman (ed.), *Elements of Grammar*, 281–337. Dordrecht: Kluwer.

Rizzi, L. 2005. On some properties of subjects and topics. In Laura Brugè, Giuliana Giusti, Nicola Munaro, Walter Schweikert & Giuseppina Turano (eds.), *Contributions to Incontro di Grammatica Generativa, Venice: Ca' Foscari University*. 203–224.

Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.

Roberts, Ian. 2014. Taraldsen's generalization and language change: Two ways to lose null subjects. In Peter Svenonius (ed.), *Functional Structure from Tip to Toe. The Cartography of Syntactic Structures* 9, 115–147. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Parameters in minimalist theory. In Teresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Samek-Lodovici, Vieri. 1996. Constraints on subjects: An optimality-theoretic analysis. PhD dissertation. Rutgers University, New Brunswick, NJ.

Shin, Naomi, L. & Caims, Helen S. 2009. Subject pronouns in child Spanish and the continuity of reference. In Joseph Collentine & Maryellen García, Barbara Lafford, Francisco Marcos Marín (eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, 155–164. Somerville: Cascadilla Proceedings.

Sobin, Nicholas. 1987. The variable status of Comp-trace phenomena. *Natural Language and Linguistic Theory* 5. 33–60.

Sorace, Antonella. 2005. Selective optionality in language development. In Leonie Cornips & Karin P. Corrigan (eds.), *Syntax and Variation*, 55–80. Amsterdam: John Benjamins.

Taraldsen, Tarald. 1978. On the NIC, vacuous application, and the that-t filter. Ms., MIT.

Vanrell, Maria M. & Fernández-Soriano, Olga 2017. Language contact at the prosody-syntax interface: Focus in Basque Spanish. Talk presented at GALA 13, 7–9 September, 2017, Palma, Spain.

Weiss, Helmut 2013. UG und syntaktische (Mikro-)Variation. [UG and syntactic (micro-) variation]. In Werner Abraham & Elisabeth Leiss (eds.), *Dialektologie in neuem Gewand. Zu Mikro-/Varietätenlinguistik, Sprachenvergleich und Universalgrammatik* [Linguistische Berichte, Sonderheft 19], 171–205. Hamburgh: Buske.

Zubizarreta, Maria-Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Axel Holvoet and Anna Daugavet

7 Types of null arguments in Baltic

1 Introduction

In this chapter we will focus on how the data of the two Baltic languages – Lithuanian and Latvian – can contribute to the discussion on generic null subjects.[1] With regard to the cross-linguistic variation observed in this domain, the position of the Baltic languages can be characterized as follows.

Both living Baltic languages, Lithuanian and Latvian, are partial Null Subject Languages (NSLs), that is, they do not require the subject position to be occupied by a phonetically non-empty pronominal element. However, in the case of anaphoric reference, the use of non-empty pronouns is frequent, though not obligatory.

3SG and 3PL generic subject constructions can be identified cross-linguistically with clearly discernible distinctive semantic features. At the same time, there is considerable variation concerning (i) the presence or absence of phonetic realization of the subject, (ii) the morphological marking on the verb, and (iii) the instantiation of the constructions, i.e. whether a language has 3SG or 3PL constructions, or both.

Though both Baltic languages have rich verbal morphology in general, 3rd person finite verb forms do not distinguish number. Other morphological features, i.e. agreement of nominal forms in case and gender with the null subject, allow null subject constructions with 3SG and 3PL to be set apart (details will be provided in Sections 2 and 3), but it is important to note that there is not always an overt morphological marker facilitating the identification of the construction involved.

Partial NSLs have been claimed to have generic 3SG null subject pronouns, as distinct from consistent NSLs, which allegedly do not have them (Holmberg 2010a: 93–94). Of the two Baltic languages only Latvian would confirm this claim. In addition, Latvian also has arbitrary 3PL null subjects, which makes this Baltic language a representative of the type exemplified by Brazilian Portuguese (see Holmberg, Nayudu & Sheehan 2009: 64). As far as we are aware, the interaction between the two different null subjects in one language has never received enough attention in literature, and our chapter aims to fill this gap.

Lithuanian, also a partial NSL like its close cognate Latvian, only has the arbitrary 3PL null subject. The issue is further complicated by the lack of number distinction in the 3rd person verbal forms in both Baltic languages.

Although Lithuanian lacks the generic 3SG null subjects and has to resort to other strategies instead, the arbitrary 3PL null subjects in Lithuanian and Latvian convey the same range of meanings. We will show that, in spite of being traditionally labelled as "arbitrary", they most frequently serve to refer to groups of people, while their other meanings can be seen as extensions of the original "corporate" meaning.

The structure of the chapter is as follows: after a section providing information on verbal morphology, indispensable for the correct understanding of the Baltic constructions under discussion and their formal identification, we will first discuss the generic 3SG null subject construction characteristic only of Latvian, and then the arbitrary 3PL null subject construction attested in both languages. In the concluding section we will address the question of cross-linguistic variation and interdependencies between the parameters of variation, in order to see what the Baltic data can contribute to current discussions.

2 Verbal morphology

The Baltic languages, Lithuanian and Latvian, have rich inflectional systems. Nouns distinguish seven cases, two numbers and two genders for agreeing adjectivals. Verbs distinguish three absolute tenses and additionally three anterior tenses composed of the auxiliary *be* and a past active participle agreeing with the subject in number and gender. Verbs also inflect for person and number, but 3rd person verb forms are systematically identical in singular and plural, see a typical conjugation of a Latvian verb *darīt* 'do' in the simple present tense in Table 1.

Table 1: The Latvian verb *darīt* 'do' in the simple present tense.

	SG	PL
1	*daru*	*darām*
2	*dari*	*darāt*
3	*dara*	

This last fact has important consequences for the subject-matter of the present article, where the distinction between 3SG and 3PL null subjects relies on agreement features. Finite verb forms cannot show number:

(1) *Mēs   esam   pilsētnieki,   kam   patīk,   ka   _____*
     we   be.PRS.1PL citydweller.NOM.PL who.DAT please.PRS.3 that (they)
     *viņus **sauc**   par   jaunajiem   lauciniekiem.*
     them   call.PRS.3 for   new.DAT.PL   countryman.DAT.PL
     'We are city dwellers who are happy to be called the new country folk.'[2]
     (Latvian)

The null subject can, however, be identified as singular or plural when the verbal form contains an agreeing participle, as in analytic perfect forms of the Latvian verb *darīt* 'do', see Table 2.

Table 2: The Latvian verb *darīt* 'do' in the present perfect tense (masculine).

	SG	PL
1	*esmu darījis*	*esam darījuši*
2	*esi darījis*	*esat darījuši*
3	*ir darījis*	*ir darījuši*

Example (2), for instance, contains a perfect form, which usually consists of a past active participle (PTCP_A) and the auxiliary 'be'. The latter can be omitted in the 3rd person, and actually is in this case. All that remains is the participle, which, by its agreement features, enables the identification of the null subject as morphosyntactically 3PL.

(2) *Tieslietu   ministrs: pēdējā   laikā   _____   mani*
     legal.affairs.GEN.PL minister latest.LOC.SG time.LOC.SG (they) me.ACC
     *(**ir**)   bieži   **saukuši**   par   reformatoru.*
     be.PRS.3 often   call.PTCP_A.NOM.PL.M for reformer.ACC.SG
     'The minister of justice: Lately they have often been calling me a reformer.'[3]
     (Latvian)

Apart from participles occurring in the perfect and evidential forms of the verb, agreeing predicative adjectivals also give us a clue as to the morphosyntactic features of the null subject. An instance of this may be seen in the following example:

2 https://www.db.lv/zinas/mana-pieredze-velas-attistit-medus-konfeksu-razosanas-liniju-492631.
3 https://juristavards.lv/zinas/263637-tieslietu-ministrs-pedeja-laika-mani-biezi-saukusi-par-reformatoru/

(3) [*Turklāt viņš uzskata*]
     moreover he thinks
     ka   drošs   _____ var   būt   tikai par
     that sure.NOM.SG.M (one) be.able.PRS.3 be.INF only   about
     to,   ar   ko   kādreiz ir   strādājis.
     them with who.ACC once   be.PRS.3 work.PTCP_A.NOM.SG.M
     '[Moreover he thinks] one can be sure only about those with whom one has
     worked at some point.'
     (Latvian, lvTenTen14)[4]

Here the nominative singular masculine adjective *drošs* as well as the perfect in
the relative clause (which contains the singular masculine participle *strādājis*)
enable the identification of the type of null subject involved.

We will show further how these two null subjects, which are only partly dif-
ferentiated by verbal agreement, are used each in its respective constructions.

3 The constructions involved

With the background information in the previous section in mind we can now give
a brief survey of the Baltic constructions we will be referring to in this chapter.

3.1 Arbitrary 3PL null subjects

The construction with an arbitrary 3PL null subject is common to both Baltic
languages. It has an exact counterpart in the Slavic languages (Lindseth 1998)
as well as in Romance consistent NSLs (Cabredo-Hofherr 2003), and a rough
equivalent in constructions with a non-referential overt 3PL pronoun like English
they say. We give a Lithuanian example with a verb that enables identification
of the null subject as 3PL. It is in the evidential mood form, which consists in
the use of an agreeing participle instead of a finite verb form. Additionally, this
example contains a converb controlled by the null subject and showing number
and gender agreement features, as shown in (4).

---

4 lvTenTen14 is the Latvian Web Corpus.

(4) *Ligą* _____ **galėję** *nuplauti,*
disease.ACC.SG (they) be.able.PTCP_A.NOM.PL.M wash.away.INF
maudydamiesi *saulėtekyje ar saulėlydyje, per Velykas ar*
bathe.CVB.NOM.PL sunrise.LOC or sunset.LOC on Easter or
Jonines.
St John's.day
'People were [reportedly] able to wash away a disease by bathing at sunrise
or sunset, on Easter or Midsummer.'
(Lithuanian, WaC)[5]

3.2 3SG null subjects

The construction with a 3SG null subject occurs only in Latvian, and it has a coun-
terpart in the Finnish "missing person construction" (Hakulinen & Karttunen
1973) or "zero person construction" (Jokela & Plado 2015), which is also known
from Holmberg (2005, 2010b). The subject is 3SG generic and corresponds to
English *one*. Agreeing forms are in the masculine singular form, (5).

(5) *Agrāk* _____ *to* *tā* **nemanījis,** *bet tagad*
formerly (one) that.ACC so NEG.notice.PTCP_A.NOM.SG.M but now
kravas *mašīnu* *kļuvis* *nenormāli* *daudz.*
charge.GEN car.GEN.PL become.PTCP_A.NOM.SG.M exceedingly many.
'Formerly [she said] one wouldn't have noticed it, but nowadays the number
of lorries has grown beyond measure.'[6]
(Latvian)

Like the 3PL arbitrary null subject discussed in 3.1, the 3SG generic null subject is
able to control converbs. These then appear in the NOM.SG form (6).

(6) *Jā,* *atceros.* *Tādus* *murgus* *pat*
yes remember.PRS.1SG such.ACC.PL nightmare.ACC.PL even
gribēdams _____ **nevar** *aizmirst.*
want.CVB.NOM.SG (one) NEG.be.able.PRS.3 forget.INF
'Yes, I remember. Such a nightmare one cannot forget, even if one tries very
hard.'
[Latvian, lvTenTen14]

---

5 LithuanianWaC is the Lithuanian Web Corpus.

6 https://www.apollo.lv/4762322/jaunciema-gatve-nesakartota-un-bistama

The absence of this construction in Lithuanian and its presence in Finnish suggest that this feature might be a Latvian-Fennic areal convergence (as suggested in Holvoet 2001).

3.3 2sg generic null subjects

As in many other languages, the 2sg form can have a generic function. In Lithuanian it may correspond to the Latvian generic 3sg null subject, see (7) from a parallel corpus:

(7) a. *Kad* _____ ***raksta*** *katru* *dienu,*
   when (one) write.PRS.3 every.ACC day.ACC
   (*nav iespējams vismaz reizi gadā nepasacīt kaut ko vērtīgu*).
   'When one writes every day, (it is impossible to avoid saying something of worth at least once a year).'
   (Latvian, [LiLa][7])

 b. *Kai* _____ ***rašai*** *kasdien,*
   when (you) write.PRS.2SG every.day
   (*neįmanoma sykį per metus nepasakyti ką nors vertinga*).
   'When one writes every day,
   (it is impossible to avoid saying something of worth at least once a year.)'
   (Lithuanian)

In Latvian, a generic 2sg pronoun may refer back to a 3rd person null subject, as can be seen from (8). This example also shows that the generic 'you' pronoun need not be phonetically empty:

(8) *Kā* *lai* _____ ***cīnās*** *ar* *to,* *ko* *tu*
   how HORT (one) fight.PRS.3 with that what 2SG.NOM
   nepazīsti?
   NEG.know.PRS.2SG
   'How is one to struggle with something that you don't know?'
   (Latvian, lvTenTen14)

---

7 LiLa is the Lithuanian-Latvian Parallel Corpus.

Generic 'you' is never silent if it is not the subject, as shown in (9).[8]

(9) *Kā   lai   ____   **saprot**,   ko   no   **tevis***
     how HORT (one) understand.PRS.3 what.ACC from 2SG.GEN
     gaida?
     expect.PRS.3
     'How is one to understand what people expect from you?'
     (Latvian, lvTenTen14)

3.4 The impersonal passive

Another expression functionally similar to that with the 3rd person null subject is provided by non-agreeing passive participles (PTCP_p.NA), also known as "impersonal passives", which are found in both Baltic languages. As in all languages that have impersonal passives, they refer to events involving humans. Impersonal passives are quite productive in Lithuanian, more so than in Latvian; parallel corpora abound in examples where a Latvian 3PL null subject construction corresponds to a Lithuanian impersonal passive, as shown in (10).

(10) a. *Par   to   ____   jau   raksta   un   runā.*
          about that.ACC (they) already write.PRS.3 and talk.PRS.3
          'People are already writing and talking about that.'
          (Latvian, LiLa)
      b. *Apie   tai   jau   **rašoma**   ir   **kalbama**.*
          About that.ACC already write.PTCP_p.NA and talk.PTCP_p.NA
          'People are already writing and talking about that.'
          (Lithuanian)

The impersonal passive has a human agentive argument in its semantic structure, but it is not clear whether it is represented in syntactic structure. For analogous Finnish constructions this assumption is made in Holmberg (2005, 2010b), but we regard it as somewhat speculative as the control test used in Sections 3.1 and 3.2 to prove the presence of null subjects yields negative results here. A converb can refer to the semantic agent argument, but in this case, it will be the non-agreeing converb in *-ant/-int*, which cannot be controlled by a syntactic subject NP, (11).

---

8 The second 3rd person form in this Latvian example represents an instance of the arbitrary 3PL null subject.

(11) *Spaudžiant   nuorodą,   šokama   į   temos   pirmąjį   puslapį.*
     press.CVB   link.ACC   jump.PTCP_PN   to   theme.GEN   first.ACC   page.ACC
     'By clicking on this link, one jumps to the first page of the [given] topic.'
     (Lithuanian, WaC)

Even in Lithuanian this form is, however, most frequent in the written language, whereas the spontaneous spoken language prefers the 3PL construction. In what follows we will concentrate on the 3PL constructions.

In this chapter, we mostly concentrate on the constructions introduced in Sections 3.1 and 3.2, where a subject, though without phonological realization, is demonstrably present in the syntactic structure (its presence is revealed by its ability to control converbs). More generally, Baltic languages can be classified as partial NSLs, in which "generic pronouns can and must be null" (Holmberg 2005: 540; Roberts & Holmberg 2010: 11–12); this property is illustrated in Latvian examples like (5) and (6). Another feature, characteristic of this type, is that referential subjects with antecedents may appear either as overt pronouns or as null pronouns (*pro*); see (12a) and (12b), respectively.

(12) a. *Į   mus   kreipiasi   įvairūs   žmonės_i   ir   jie_i*
          to   us   apply.PRS.3   various.NOM.PL   people.NOM.PL   and   they
          žinoma   nenurodo   savo   tautybės.
          of_course   NEG.mark.PRS.3   their   nationality.GEN
          'Various people apply to us and they, of course, do not mark their nationality.'
          (Lithuanian, WaC)

      b. *Ar   vaikai_i   irgi   kalti,   kalti,   kad   _____i*
          Q   child.NOM.PL   also   guilty.NOM.PL   guilty.NOM.PL   that   (they)
          turi   tokius   "kvailus"   tėvus?
          have.PRS.3   such.ACC.PL   stupid.ACC.PL   parent.ACC.PL
          'Is it the children's fault as well – is it their fault they have such "stupid" parents?'
          (Lithuanian, WaC)

The factors determining the distribution of overt and null referential subjects are complex and will not be discussed here. Suffice it to say here that, unlike what we observe in consistent NSLs, the use of an overt pronoun is not a marked option: it does not necessarily mark a change in discourse referent.

4 The Latvian generic 3SG null subject construction

As in the Baltic languages both overt and null pronouns may have antecedents, a verb form without an overt subject may have, in principle, two readings, an anaphoric/definite and a non-anaphoric/indefinite one. As we will see, the details are slightly different in the case of 3SG and 3PL null subjects.

As mentioned in Section 3.4, generic 3SG null subjects are attested only in Latvian, not in Lithuanian. Their meaning is generic inclusive, whereas arbitrary 3PL null subjects, to the extent that they are at all generic, are almost always generic exclusive (Holmberg 2010a: 92–93). It is not the case that 3SG forms are consistently ambiguous between individual and generic meaning, the syntactic context deciding which interpretation applies. The generic readings are, in fact, subject to rigid constraints. The elements licensing a generic reading of the 3SG verb form are of two kinds: lexical and syntactic. We will start with the lexical licensing factors.

4.1 Modal verbs

Modal verbs are the lexical items most regularly allowing a generic meaning: most frequently *var* 'one can', less frequently *drīkst* 'one may'. To illustrate our point, we have selected examples with agreeing predicate adjectives, revealing the 3SG.M agreement features, wherever possible:

(13) *Satiksmē*   *(arī* *uz* *ietves)*   ***nedrošs***   _____
      traffic.LOC also   on   pavement.GEN insecure.NOM.SG.M (one)
      nedrīkst   *justies.*
      NEG.be.allowed.PRS.3   feel.INF
      'In the traffic, including the pavement, one is not supposed to feel insecure.'
      (Latvian, lvTenTen14)

4.2 Verbs of perception

A second lexical group regularly allowing a generic reading is formed by the verbs of perception: *redz* 'one can see', *dzird* 'one can hear', *jūt* 'one can feel', as illustrated in (14).

(14) *Ezers   ir   tik rāms,   ka   _____   **neredz***
lake.NOM.SG be.PRS.3 so quiet.NOM.SG.M that (one)   NEG.see.PRS.3
nevienu,   pat   ne   vismazāko,   vilnīti.
no.ACC.SG even NEG smallest.ACC.SG wave.ACC.SG
'The lake is so quiet that one does not see even the smallest ripple.'
(Latvian, lvTenTen14)

4.3 Mood

There are several syntactic contexts that may license a generic reading of the 3SG null subject. The most important is conditional mood. A generic null subject may occur both in the protasis and in the apodosis, and the two are then consistently interpreted as coreferential. Under these conditions, any verb (not just a verb belonging to the lexical classes mentioned in sections 4.1 and 4.2) may receive a generic interpretation. None of the lexical licensing factors need be present in such cases:

(15) *Ja   _____   no rīta   izkāpj   no   gultas   ar*
if   (one) in the.morning step.out.PRS.3 from bed.GEN   with
kreiso   kāju,   tad   _____   būs   dusmīgs   un
left.ACC leg.ACC then (one)   be.FUT.3 angry.NOM.SG.M and
saīdzis   visu   dienu.
peevish.NOM.SG.M whole.ACC day.ACC
'If one steps out of one's bed with one's left leg, then one will be angry and peevish all day.'
(Latvian, Straubergs, *Latviešu tautas paražas*)

4.4 Purpose clauses

Another preferential context for 3SG null subjects is purpose clauses with coreferential subjects in main and subordinate clause. The subordinate clause obligatorily contains an irrealis form. Although in modern Latvian the irrealis does not inflect for person and number any more, the verbal form can with reasonable certainty be identified as 3rd person (the only alternative would be 2SG, but then

we would expect 2sg subject pronouns to surface now and then), and the M.SG value can be seen from the agreeing adjectivals, as in (16).

(16) *Lai* _____ *būtu* *draudzīgs* *pret* *vidi,*
      in.order.to (one) be.IRR friendly.NOM.SG.M toward environment.ACC
      pirmkārt *ir* *jāšķiro* *atkritumi.*
      first.of.all be.PRS.3 DEB.sort waste.NOM
      'In order to be environment-friendly, one should first of all sort household waste.'
      (Latvian, lvTenTen14)

This could be paraphrased as: 'for any person to be environment-friendly, it is necessary etc.', which is in line with the reading of 3sg null subjects in other syntactic contexts.

4.5 Generic null subjects in questions

Another characteristic type of generic null subjects is found in deontic (deliberative) questions (i.e. questions meant to elicit a directive rather than a piece of information). When the (often rhetorically) solicited directive refers to a course of action to be taken by a generic referent, a non-finite form, like the infinitive, is often used (*Where to go now?*). In Latvian, such deontic questions usually contain the hortative particle *lai* with a generic 3sg null subject, (17).

(17) *Kā* *lai* _____ **kļūst** **vesels** *bez*
      how HORT (one) become.PRS.3 healthy.NOM.SG.M without
      kādām *zālēm?*
      any.DAT.PL medicine.DAT.PL
      'How to become healthy without any medicines?'
      (Latvian, lvTenTen14)

This type is also interesting in that, under embedding, it sometimes seems to lose its generic meaning and to be controlled by an individual, referential main clause subject of any person value, as in (18).

(18) *Pirmo reizi dzīvē man kaut kas tik ļoti pietrūkst,*
first time life.LOC I.DAT something so much be.lacking.PRS.3
ka es tiešām nezinu, ko _____ lai dara.
that I really NEG.know.PRS.1SG what.ACC (one) HORT do.PRS.3
'For the first time in my life I so strongly feel the lack of something that I really don't know what to do.'[9]
(Latvian)

This could be compared to a similar construction with the infinitive:

(19) *Ja tiešām nezini, ko* PRO *dāvināt puisim*
if really NEG.know.PRS.2SG what.ACC gift.INF boy.DAT
Ziemassvētkos, [*iedomājies par viņa interesēm un hobijiem.*]
Christmas.LOC imagine.IMP.2SG about his interests and hobbies
'If you really don't know what to give to a boy for Christmas, [think about his interests and hobbies.'][10]
(Latvian)

It seems that the null subject of the 3SG form behaves, in such contexts, as a PRO subject controlled by the main clause subject.

With the exception of contexts involving the lexical licenser mentioned in section 4.2 and the modal licensing in examples like (13), the construction with 3SG null subjects does not occur in simple clauses. If we omit the modal verb from (3), the sentence becomes ungrammatical if the subject is to be interpreted as generic:

(20) **Drošs _____ ir tikai par to, ar ko*
sure.NOM.SG.M (one) be.PRS.3 only about them with who.ACC
kādreiz ir strādājis.
once be.PRS.3 work.PTCP$_A$.NOM.SG.M
'One is sure only about those with whom one has worked.'
(Latvian)

In order to be grammatical, this sentence must have a contextually retrievable antecedent ('He is sure only about those with whom he has worked').

---

9 https://laumastankevica.tumblr.com/page/2
10 https://www.fromme.lv/raksti/puisim-davana-ziemassvetkos.html

4.6 Typology of 3SG null subjects

Similar sets of licensing factors can be observed in other languages that have generic 3sg null subjects. For Icelandic, Sigurðsson & Egerland (2009: 160) mention the "impersonal modal construction", where a modal verb licenses a 3sg null subject, whereas the corresponding sentence without modal verb is ungrammatical, as shown in (21).

(21) a. *Hér   má   ekki dansa.*
       here may.PRS.3SG NEG dance.INF
       'One is not allowed to dance here.'
       (Icelandic, Sigurðsson & Egerland 2009: 167)

   b. **Hér   dansar   oft.*
       here dance.PRS.3SG often
       'One dances often here.' (intended)
       (Icelandic, Sigurðsson & Egerland 2009: 167)

In a similar vein, Middle Polish shows two contexts licensing 3sg null subjects, i.e. modal verbs (22a) and conditionals (22b).[11, 12]

(22) a. *W Przewodni tydzień śnieg   i   mróz wielki, że*   _____
       in Easter   week   snow and frost big   that (one)
       mógł   saniami jechać.
       could.M.SG   sleigh.INS ride.INF
       'During Easter Week there was snow and such a big frost that one could ride a sleigh.'
       (Middle Polish, Jan Chryzostom Pasek, 17th century)

---

11 The presence of a generic 3sg null subject in Middle Polish is unexpected as Polish is a typical representative of consistent NSLs, which are not supposed to have generic 3sg null subjects (Holmberg 2010a: 93).

12 In Middle Polish the class of perception verbs also included *najść* 'find' and that of modal verbs included *dostać* '(manage to) get'. On the 3sg null subject in the history of Polish, see Pisarkowa (1984: 30–31).

b. *A   tam   przysmak   taki,   że=by   koza   wrzeszczała,*
and   there pungent.taste such   that=IRR goat.NOM screamed.F.SG
gdy=by _____ *jej   gwałtem   wlał.*
if=IRR (one) her.DAT   forcibly   poured.into.M.SG
'[The taste (of the liquor) was so pungent] that a goat would have
screamed if one had forced it down its [throat].'
(Middle Polish, Jan Chryzostom Pasek, 17th century)

The widest set of contexts allowing 3SG null subjects that we are aware of is
found in Finnish, but here as well there has to be a licenser. In addition to the
modal verbs *voi* 'can', *täytyy* 'must' and *saa* 'may', Hakulinen & Karttunen (1973:
160–161) mention a list of verbs which they call broadly "modal", but which also
includes verbs that would have to be regarded as implicatives in the sense of
Karttunen (1971), such as *uskaltaa* 'dares', *viitsii* 'feels like' or *tarkenee* 'is warm
enough'. Implicatives are a type of complement-taking predicates related to
modals (especially to 'dynamic' modals, on which see Palmer 2001: 9–10, 76–80).
Whereas modals evaluate potential events or event types in terms of possibility or
necessity, implicatives characterize mental, physical or emotive states as neces-
sary and sufficient conditions for the accomplishment of an event. The factual or
counterfactual implications constituting the definitional feature of implicatives
do not hold when the sentence is generic. We can thus characterize modals and
implicatives as operating on potential events, and in this sense constituting a
diminished referentiality context.

Finnish allows one more type of licenser, namely adverbials. A characteristic
type is that of adverbials of manner, as in (23).

(23)   *Sen   työ   ___   tekee   helposti.*
this.GEN.SG job.nGEN.SG (one) do.PRS.3SG easily
'That job you can easily do.'
(Finnish, Hakulinen & Karttunen 1973: 163)

Such sentences come very close to so-called "middle-voice constructions", as in
English *This book reads well*, or German *Es wohnt sich hier gut* 'It's good to live
here'. Middles typically imply a generic agent, in many European languages they
have a morphological marker (typically reflexive), and they are usually accompa-
nied by a manner verb like *well*, *easily*, *comfortably* etc. indicating that external
circumstances rather than the subject's agency are responsible for the success-
ful realization of a state, activity, process etc. The manner adverbial in (23) thus
acts as a licenser of generic interpretation because it points to the middle-voice
reading.

The set of licensing factors for generic readings of 3SG null subjects may thus be larger or smaller, but it has a constant nucleus to which modals (as a lexical licenser) and hypothetical/conditional context (as a syntactic licenser) certainly belong. There might be a hierarchy of licensing factors, but it is not easy to establish because languages may show occasional gaps. In conditionals, for instance, Finnish has 3SG null subjects in generic function but Estonian uses the infinitive instead (Jokela & Plado 2015). In Middle Polish, lexemes licensing 3SG null subjects include modal verbs but not verbs of perception, as Polish has specialized modal perception verb constructions such as *widać* 'see.INF', which have a generic experiencer in their semantic structure but no subject position in their syntactic structure. Thus, while the set is more or less constant, a specific type of null subject licenser may be lacking in a particular language because an alternative strategy is available that makes the null subject construction redundant.

5 The arbitrary 3PL null subject construction

While generic 3SG null subject constructions are characteristic only of Latvian, arbitrary 3PL null subject constructions are used in both Latvian and Lithuanian. Although, as mentioned in Section 2, there are no verbal forms distinguishing 3SG and 3PL, the two constructions are well kept apart in Latvian, where they coexist, and they don't seem to overlap, though similarities exist.

As in other languages that have arbitrary 3PL null subjects, those of Baltic may have a range of interpretations. The following classification has been proposed by Cabredo-Hofherr (2003), which is adopted with only minor modifications by Siewierska & Papastathi (2011). The distinguished types are universal, corporate, inferred existential, vague existential and specific existential. In what follows we will basically be using the terms proposed by Cabredo-Hofherr; however, we will not invoke the distinction between vague existential and specific existential, though we do not want to deny the relevance of the distinction. Cabredo-Hofherr's classification consists of more or less discrete types, but the different readings could also be conceived as constituting a cline ranging from universal to vague-individualized. A similar picture emerges for Baltic: unlike the Latvian 3SG null subject, which has only one reading, arbitrary 3PL null subjects have a number of readings in both languages. In line with Siewierska & Papastathi (2011), we will be concerned with diachronic relationships between these types as determining their patterns of instantiation. Since Siewierska & Papastathi's results are partly based on the comparison between translations of the first book of the Harry Potter series into several European languages (with the

notable exception of Baltic), we will illustrate our points with examples from the Latvian and Lithuanian versions of *Harry Potter and the Philosopher's Stone* and its sequel *Harry Potter and the Chamber of Secrets*, which were not included in Siewierska & Papastathi's research.

One of the most important conclusions reached by Siewierska & Papastathi (2011) is that many languages only marginally accept inferential and vague readings, whereas there are generally no objections to corporate and universal readings. There must be good reasons for this, and we assume they are basically diachronic: the point of departure for arbitrary 3PL null subjects must be the anaphoric or deictic use of 3PL forms (or, in the case of Baltic, undifferentiated 3SG/PL forms) with reference to groups of persons. At a certain point, 3PL constructions may establish themselves as a means of referring to a group of persons that is not anaphorically or deictically given but may be identified in other ways. At a next stage, restricted to a subset of languages with 3PL null subjects, these constructions may extend to refer to individual unidentified subjects.

5.1 Corporate and universal interpretations of 3PL null subjects

It is important to pause over the mechanisms of the extensions involved in this process. The prominence of corporate and universal readings is natural as both types of situation provide for an easy way of identifying the referents on the basis of general knowledge: one identifies the people who could be involved in a certain type of event in virtue of their social function (corporate) or one generalizes 'they' to the members of the community of which one is part (universal). It is not clear whether either of these two is more basic than the other; Siewierska & Papastathi (2011) find no difference in distribution (languages having one also have the other), but they hypothesize a development *known* (anaphoric) → *partially known explicit* (universal) → *partially known deduced* (corporate). In the absence of empirical evidence (in the form of a language having the universal but not the corporate type) this remains speculative, and one could probably also make a case for the reverse development – the universal type as an extension from the corporate type.[13] But as this discussion is devoid of an empirical foundation and not immediately relevant to our main topic, we will not pursue it.

---

13 Siewierska & Papastathi's (2011) hypothesis is based on the claim that universal uses have to be licensed by locative phrases (*In Spain they speak Spanish*). But this claim does not seem to have been subjected to empirical verification, and whether what has been claimed for Spanish and English can be upheld, for instance, for Slavic and Baltic, is unknown.

As other authors have done before us, we will use the notion of *corporate* in a broader sense than Pesetsky's (1995), who uses it to denote "[a] pronoun [that] picks out some socially designated group of people, prototypically governments, bosses, criminals, or shopkeepers", which suggests that the group involved must be socially established and readily recognizable within a certain community. But the notion should be more flexible so as to accommodate categories construed *ad hoc* in specific situations and thus requiring a lot of contextual knowledge, as argued by Siewierska & Papastathi (2011: 592). In the uncanny-boarding-school setting of the Harry Potter translation corpus used by Siewierska & Papastathi, 'they' may refer to the teaching staff, but also to the pupils collectively, the supernatural powers at work in the school, and so on. There is thus no finite list of socially recognizable groups, and no finite list of event types typically performed by them, to define the corporate use. The corporate type also happens to be the most frequent one among all indefinite uses of the 3PL null subjects that we were able to find in the first two Harry Potter volumes.[14]

(24) a. *He was going to be expelled, he just knew it.*
      b. *Tagad* ____ *viņu izslēgs,   par   to*
         now   (they) him expel.FUT.3 about that.ACC
         nevarēja   būt   ne   mazāko   šaubu.
         NEG.be.possible.PST.3 be.INF even slightest doubt.GEN.PL
         Lit. 'They would now expel him, there could be little doubt about that.'
         (Latvian)

5.2 Individualized (vague) interpretations of 3PL null subjects

The shift from group reference to individualized (vague) reference may be assumed to involve the mechanism of representation: an institution may be represented, in specific instances, by one of its members, as in the case of a postman delivering the mail. The shift is then often twofold: from collective to individual representative, and from habitual (event type) to temporally anchored (event token). We will refer to this as the corporate-instantiated use.

(25) a. *Good wand, that one. But I suppose they snapped it in half when you got expelled?*

---

14 We give the original English sentences corresponding to the Latvian and Lithuanian translations, but where these diverge too strongly from the original, a separate literal translation is added.

b. *Tas   bija   labs   zizlis.   Bet,   domājams,   _____   to*
that   was   good   wand   but   presumably   (they)   it.ACC
salauza,   kad   _____   jūs   izslēdza   no
break.PST.3   when   (they)   you.ACC   expel.PST.3   from
skolas?
school.GEN.SG
Lit. 'It was a good wand. But, presumably, they broke it when they expelled you from the school?'
(Latvian)

c. *Gera   buvo   lazdelė.   Tačiau,   tikiuosi,   _____   ją*
good   was   wand   however   believe.PST.1SG   (they)   it
sulaužė,   kai   _____   tave   išmetė   iš
break.PST.3   when   (they)   you.ACC   expel.PST.3   from
mokyklos?
school.GEN.SG
Lit. 'It was a good wand. But I hope they broke it when they threw you out of the school?'
(Lithuanian)

It presumably took only one person to snap the wand in two though it was done on behalf of a group – in this case, the Ministry of Magic.

A characteristic feature of the corporate-instantiated use is that the group is identifiable but the individual within the group remains unidentified and its identity is viewed as irrelevant. This type can, in its turn, give rise to two extensions: in one, an individual remains unidentified and no group anchoring it in discourse space is available to the hearer, as is the case with corporate use:

(26) a. *The body of the girl who was killed was discovered in a bathroom.*
b. *Nogalināto   meiteni   _____   atrada   tualetē.*
kill.PTCP$_{p}$.ACC   girl.ACC   (they)   find.PST.3   toilet.LOC
Lit. 'They found the killed girl in a bathroom.'
(Latvian)

The other extension leads to uses where the referent is situationally retrievable, just as the group is retrievable in the corporate type, but (s)he is not explicitly identified. This lack of explicit identification is enabled by the fact that it is backgrounded. This is seen in (27), where the null subject of the Latvian *ved* / Lithuanian *atvedė* is not anaphoric; though known from the broader context, the subject referent is too remote to be accessible for anaphoric reference, but, being backgrounded, it can loosely be referred to by a 3PL null subject:

(27) a. [*Harry could see a gleaming oak door ahead, with a brass knocker in the shape of a griffon.*] *He knew where he was being taken.*

b. [. . .] *Viņš saprata, kurp _____ viņu **ved**.*
   he understand.PST.3 whither (they) him lead.PRS.3
   Lit. '[. . .] He understood where they are taking him.'
   (Latvian)

c. [. . .] *Jisai suprato, kur _____ jį **atvedė**.*
   he understand.PST.3 where (they) him bring.PST.3
   Lit. '[. . .] He understood where they are taking him.'
   (Lithuanian)

This 3PL null subject could also be called "vague", but not "vague unknown" as in (25); the referent is known but irrelevant. We could characterize this type as "vague backgrounded".

The vague backgrounded uses of the arbitrary 3PL null subject seem to be a point of departure for a final type to be discussed here, which we could call 'generic backgrounded'. The vague backgrounded 3PL in (27) occurs in a 'realis', temporally anchored clause. The character is actually being led somewhere by a person whose identity has been previously specified but is irrelevant in the given situation. But this use of the 3PL null subject in contexts where the referent is backgrounded and its identity is irrelevant can be carried over to contexts with predicates selecting non-anchored complement clauses, like 'being accustomed to', 'like', 'hate, dislike' and the like, where the complement denotes an event type rather than a temporally anchored event token. The reference of the 3PL null subject then becomes generic, as can be seen in (28).

(28) a. *Dudley wasn't used to being ignored.*

b. *Dūdijs nebija pieradis, ka par viņu*
   Dudley NEG.be.PST.3 accustomed.NOM.SG.M that about him
   _____ **neliekas** *zinis*.
   (they) NEG.pretend.PRS.3 aware
   Lit. 'Dudley wasn't used that they ignore him.'
   (Latvian)

c. *Dudlis nebuvo pratęs, kad jo*
   Dudley NEG.be.PST.3 accustomed.NOM.SG.M that him
   _____ **nepastebėtų**.
   (they) NEG.notice.IRR.3
   Lit. 'Dudley wasn't used that they wouldn't notice him.'
   (Lithuanian)

Here the 3PL null subject is truly generic in the same sense in which Latvian 3SG null subjects are generic: reference is made to a type of event involving any conceivable human subject, the speaker not excluded. The distinctive feature of almost all arbitrary 3PL null subjects, viz. that of being exclusive (i.e., not including the speaker), no longer seems to apply here. We could thus speak of a certain overlap between 3SG and 3PL null subjects in Latvian, as 3PL null subjects can also marginally be generic. But it is not clear there are instances in which 3SG and 3PL null subjects could be used interchangeably, as both function within well-defined constructions. 3PL generic null subjects seem always to be embedded and do not extend to the main sentential predication.

Among the types singled out by Cabredo-Hofherr (2003) and Siewierska & Papastathi (2011), the inferred type does not seem to be represented at all in Baltic. Inferential meaning is usually expressed by constructions with passive participles. The relationship between the English and Latvian/Lithuanian constructions is reversed: while Baltic 3PL null subject constructions often correspond to English passives (especially when the subject is backgrounded, as in (26), (27)), the Baltic passive corresponds to an English 3PL null subject construction:

(29) a. [*Your platform should be somewhere in the middle,*] *but **they** don't seem to have built it yet, do they?*

    b. [. . .] *bet, šķiet,*   *tā vēl nav*   *uzbūvēta.*
        but appear.PRS.3 it yet be.NEG.PRS.3 build.PTCPₚ.NOM
        Lit. '[. . .] but it seems it hasn't been built yet.'
        (Latvian)

    c. [. . .] *bet, matyt,*   *dar nepastatytas,*   *ką?*
        but apparently yet NEG.build.PTCPₚ.NOM what
        Lit. '[. . .] but it seems it hasn't been built yet, has it?'
        (Lithuanian)

This picture of the 3PL null subject in Baltic is based on a relatively small and undifferentiated corpus and it is probably susceptible of correction both with regard to its empirical foundation and to the way the usage types are singled out. But even with these caveats it seems we may conclude the range of usage types of 3PL null subjects as reflected in Baltic goes beyond what has already been identified in Cabredo-Hofherr (2003) and Siewierska & Papastathi (2011). In particular, the known backgrounded and the generic backgrounded types do not seem to be covered by any of the types hitherto singled out.

6 In conclusion

Both living Baltic languages are partial NSLs, but they differ with regard to their use of generic and arbitrary null subjects. Whereas Latvian has generic 3sg null subjects, these are unknown in Lithuanian, which uses generic 2sg forms instead.

The use of the generic 3sg null subject in Latvian is lexically and syntactically restricted, and the restrictions are reminiscent of those known from other languages. As in Finnic and Icelandic, the generic 3sg null subject is only found with modal verbs and verbs of perception in main clauses. This lexical restriction is however lifted in conditional period and in purpose clauses with coreferential subjects in main and subordinate clause, as well as in deliberative questions.

Although both Baltic languages have arbitrary 3pl null subjects, their use in Lithuanian is less frequent in the written variety, as impersonal or agentless personal passive forms are used instead. Nevertheless, the arbitrary 3pl null subjects in Lithuanian and Latvian appear to have the same range of meanings. Most frequently they receive a corporate reading, which is however understood more broadly than it is usually done in the literature. Apart from the information provided by the predicate and its arguments, the corporate reading is also identified with the help of contextual knowledge and general world knowledge. The prevalence of the corporate reading in the data can be taken as evidence that the arbitrary 3pl null subjects basically refer to groups of humans.

The established classification involving universal, corporate, and three subtypes of existential reading (specific, vague and inferential) can be replaced by a cline ranging from universal to vague-individual meaning. Through the corporate-instantiated meaning, which involves an institution as presented by one of its members, the non-anaphoric 3pl null subjects develop uses where they have individual referents which can be situationally retrievable or not. This development has no parallels in those Romance and other languages on which the analysis of the arbitrary 3pl null subjects has been based so far. On the other hand, inferential reading is never expressed by the arbitrary 3pl null subject in the Baltic languages, as the latter consistently employ passive forms for this purpose.

As can be seen from our discussion of null subject constructions in Baltic, the deficient number marking on the verb observed in both Baltic languages does not stand in the way of a consistent functional differentiation of generic 3sg and arbitrary 3pl null subject constructions. It is apparently the semantic features of the null subject that determine singular or plural agreement whenever it can manifest itself. Verbal morphology is not, in itself, a driving force.

This does not mean, however, that verbal morphology is irrelevant to the functioning of singular and plural null subjects. The constraints imposed on the occurrence of generic 3sg null subject constructions, contrasted with the lack of

such constraints in the case of arbitrary 3PL null subject constructions, suggest that any 3SG pronominal, whether overt or silent, is naturally biased towards an individual referential reading to a much higher extent than 3PL pronouns. This is supported by the fact that in many languages overt 3PL pronouns (or at least their non-emphatic variants) can have indefinite readings, e.g. English *they*, Dutch *ze*, whereas there are no analogous overt indefinite 3SG pronouns. The same principle evidently holds for null pronominals. With reference to Latvian, this claim may seem paradoxical considering that usually this language does not formally distinguish 3SG and 3PL verb forms. Perhaps we should say that when no contextually retrievable individual referent is available, a 3rd person null subject is by default interpreted as 3PL unless one of the licensing factors is present.

Sources

Latvian Web Corpus (lvTenTen14). https://www.sketchengine.eu [accessed 15th Dec. 2020]

Lithuanian-Latvian Parallel Corpus (LiLa). http://lila.korpuss.lv [accessed 15th Dec. 2020]

Lithuanian Web corpus (LithuanianWaC). https://www.sketchengine.eu [accessed 15th Dec. 2020]

Pasek, Jan Chryzostom. Pamiętniki [1656]. http://wolnelektury.lv [accessed 15th Dec. 2020]

Rowling, J. K. *Harry Potter and the Philosopher's Stone*. Bloomsbury, 1997. = Roulingsa, Dž. K. *Harijs Poters un filozofu akmens*. Latvian translation Ingus Josts. Jumava, 2001. = Rowling, J. K. *Haris Poteris ir išminties akmuo*. Lithuanian translation Zita Marienė. Alma littera, 2000.

Rowling, J. K. *Harry Potter and the Chamber of Secrets*. Bloomsbury, 1998. = Roulingsa, Dž. K. *Harijs Poters un noslēpumu kambaris*. Latvian translation Ingus Josts. Jumava, 2001. = Rowling, J. K. *Haris Poteris ir paslapčių kambarys*. Lithuanian translation Zita Marienė. Alma littera, 2000.

Straubergs, Kārlis. *Latviešu tautas paražas* [Latvian Popular Customs]. Rīga: Latvju grāmata, 1944.

References

Cabredo-Hofherr, Patricia. 2003. Arbitrary readings of third person plural pronominals. In Matthias Weisgerber (ed.), *Proceedings of Sinn und Bedeutung*, 81–94. University of Konstanz.

Hakulinen, Auli & Lauri Karttunen. 1973. Missing persons: On generic sentences in Finnish. In Claudia W. Corum, Thomas Smith-Start, & Ann Weiser (eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society* Vol. 9, 157–171. Chicago: Chicago Linguistic Society.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36(4). 533–564.

Holmberg, Anders. 2010a. Null subject parameters. In Theresa, Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. The null generic subject pronoun in Finnish: a case of incorporation in T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63 (1). 59–97.

Holvoet, Axel. 2001. Impersonals and passives in Baltic and Finnic. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages: Typology and Contact. Volume 2. Grammar and Typology*, 363–389. Amsterdam/Philadelphia: John Benjamins.

Jokela, Hanna & Helen Plado. 2015. Subjects under generic conditions. Implied subjects in Finnish and Estonian *if*-clauses. In Marja-Liisa Hellasvuo & Tuomas Huumo (eds.), *Subjects in Constructions – Canonical and Non-canonical*, 73–97. Amsterdam & Philadelphia: John Benjamins.

Karttunen, Lauri. 1971. Implicative verbs. *Language* 47(2). 340–358.

Lindseth, Martina. 1998. *Null Subject Properties of Slavic Languages. With Special Reference to Russian, Czech and Sorbian*. München: Verlag Otto Sagner.

Palmer, Frank Robert. 2001. *Mood and Modality*. 2nd edition. Cambridge: Cambridge University Press.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: MIT Press.

Pisarkowa, Krystyna. 1984. *Historia składni języka polskiego* [History of the syntax of Polish]. Warszawa: Wydawnictwo Polskiej Akademii Nauk.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Siewierska, Anna & Maria Papastathi. 2011. Towards a typology of third person plural impersonals. *Linguistics* 49(3). 575–610.

Sigurðsson, Halldór Ármann & Verner Egerland. 2009. Impersonal null subjects in Icelandic and elsewhere. *Studia Linguistica* 63(1). 158–185.

Urpo Nikanne
8 Interpreting null subjects in Finnish finite sentences

1 The goal of the chapter

This chapter discusses the licensing and semantic interpretation of null subjects in Finnish finite sentences that are not embedded sentences or other complex sentences in which an NP can be omitted if it is coreferential with another NP in the same complex sentence. The motivation of this restriction is that these cases form a group that has properties of its own. When the subject is lacking in a finite simple sentence, the lack of the subject is not based on the presence of an overt subject elsewhere in the complex sentence.

Because of the specific restriction of the focus, several kinds of null subjects are not discussed in this chapter even though they are very interesting as such. Given the limited scope of the chapter, several null subject types will not be discussed in this chapter.

Thus, non-finite structures with an infinitival or participial predicate are too numerous to be discussed here. Finnish has several kinds of infinitives and participles and they are involved in various, very diverse constructions.

Complex sentences are not discussed, either. In Finnish, as well as in many other languages, the subject argument may be absent in embedded finite and non-finite sentences alike. The lack of the syntactic subject in an embedded sentence requires the analysis of the whole complex sentence and the analysis of possible references within it. Thus, I leave out cases like in (1).

(1) a. *Tohtori   tiesi,   mitä   teki.*
         doctor.NOM know.PST.3SG what   do.PST.3SG
         'The doctor knew what (s)he was doing.'
     b. *Tohtori   tutki   potilaan   ja   määräsi*
         doctor.NOM examine.PST.3SG patient.ACC and   prescribe. PST.3SG
         lääkettä.
         medicine.PART
         'The doctor examined the patient and prescribed the medicine.'

In (1a), the subject argument of the embedded sentence (here: the subject of the verb *teki* 'was doing') is not present in syntax, but a coreferential NP can be found in the same complex sentence (here: the subject *tohtori* 'doctor' of the matrix verb

tiesi 'knew'). In the coordinated sentenence, (1b), the subject argument of the verb *määrätä* 'prescribe' is coreferential with the subject *tohtori* 'doctor' of the verb *tutkia* 'examine.'

Null subjects motivated by textual or dialogue-based factors fall outside the scope of this contribution. Thus, I will not go through textual subject omissions of the type exemplified in (2).

(2) *Tohtori ei   tiennyt,   mitä teki.   Sanoi   vain.*
     doctor NEG.AUX.3SG know.PTCP.PST what do.PST.3SG say.PST.3SG only
     'The doctor did not know what (s)he was doing. (S)he only said so.'

The subject argument of the verb *sanoi* 'said' in the latter sentence is omitted, but the predicate *ei tiennyt* 'did not know' of the former sentence is coreferential with the subject argument of the verb *sanoi*.

An example of a dialogue-based omission of a syntactic subject is provided below in (3).

(3) A: *Toimiiko   tämä   auto   sähköllä?*
         run.3SG.Q   this.NOM car.NOM   electricity.ADE
         'Does this car run by electricity?'
     B: *Toimii.*
         run.3SG
         'Yes, it does.'

In Finnish, the typical way to answer to a *yes/no* question is to repeat the predicate in an appropriate form without the subject. The omitted subject is understood to be coreferential with the subject of the question that it answers (in this case: *tämä auto* 'this car').

The focus is on the interpretation of the argument structure, and the theoretical background used in the analyses is the micro-modular approach of Conceptual Semantics (Nikanne 2018). The interpretation of sentences without a subject is explained as an interaction between three different levels of argument structure: 1. syntactic arguments (subject and object), 2. lexically determined arguments ("logical" subject and object), and 3. lexical conceptual structures. It is shown that empty syntactic subject is not needed for interpretation of the null subject. This requires that the system, as a whole, can link the structure to lexically determined arguments and a well-formed conceptual structure.

The outline of the chapter is as follows: In section 2, I introduce the most important cases in which Finnish does not have a subject in the finite sentence. Next, in section 3, I give the theoretical background of the analysis. Section 4

offers the analysis of the sentences introduced in Section 1. Section 5 concludes the discussion.

2 Null subjects discussed in this chapter

The finite clause types that do not have a subject discussed in this chapter are presented below:

(4) *Pro*-drop in 1st and 2nd person
Juoksen   kotiin.
run.1SG   home.ILL
'I run / I am running / I'll run home.'

(5) Passive sentences
Töistä   juostaan   kotiin.
work.ELA   run.PASS   home.ILL
'People run home from work.'

(6) Generic sentences
Kotiin   juoksee   tunnissa.
home.ILL   run.3SG   hour.INE
'It takes an hour to run home.'

(7) *Weather*-sentences
Sataa.
rain.3SG
'It is raining.'

In addition to these, there are several fixed constructions[1] that do not have a subject. An example at hand is the Instrumental Elative Construction (Nikanne 2005). In the non-causative version of this construction, there is no syntactic subject or object. The interpretation is that the NP in the elative is an instrument of a violent attack (e.g., hitting or stabbing), consider (8).

---

1 Constructions are licensed by construction-specific linking instructions between morpho-syntax, semantics, and the lexicon There are too many constructions that lack the overt subject argument to be discussed in any systematic way. However, we will discuss some constructions, as we will see that also generic sentences as well as weather sentences can also be based on constructions.

(8) a. *Joskus   tuli   puukosta.*
         sometimes   come.PST.3SG   knife.ELA
         'Sometimes one got stabbed with a knife.'
     b. *Mies   sai   Laajasalossa   puukosta*   naamaan.[2]
         man   get.PST.3SG Laajasalo.INE knife.ELA   face. ILL
         'A man was stabbed in the face in Laajasalo.'

3 Theoretical background

3.1 Organization of grammar

The analysis of null subjects in this article is based on the micro-modular theory of Conceptual Semantics developed in Nikanne (2018).[3] The overall organization of grammar is as presented in Figure 1.

The system is based on a network of representational modules (on representational modularity, see Jackendoff 1990). The formation rules of the representations are as simple as possible, and each representational module has its own basic categories and the rules that govern their combinations.

As the formation of the representations is very simple, the links between the representational units play a very important role. In addition to the representational modules, there are three modules that can license links between more than two representational modules. These are listed below:

1. **The Lexicon** can link fragments of phonological, syntactic, and conceptual representations with each other. In addition, lexical entries can be associated with other kinds of information.
2. **Constructions** can specify and allow links between fragments of the syntactic structure with the specified morphological categories and lexical items even when those links do not follow the productive principles of grammar.
3. **Morphology** can link fragments of phonological, lexical, and syntactic representations with each other.

---

2 Headline in *Helsingin Sanomat* newspaper 9 October 2020 (p. A 24). Laajasalo is a suburb in Helsinki.
3 On the micro-modular model, see also Nikanne (1990); Pörn (2004); Petrova (2011) and Paulsen (2011).
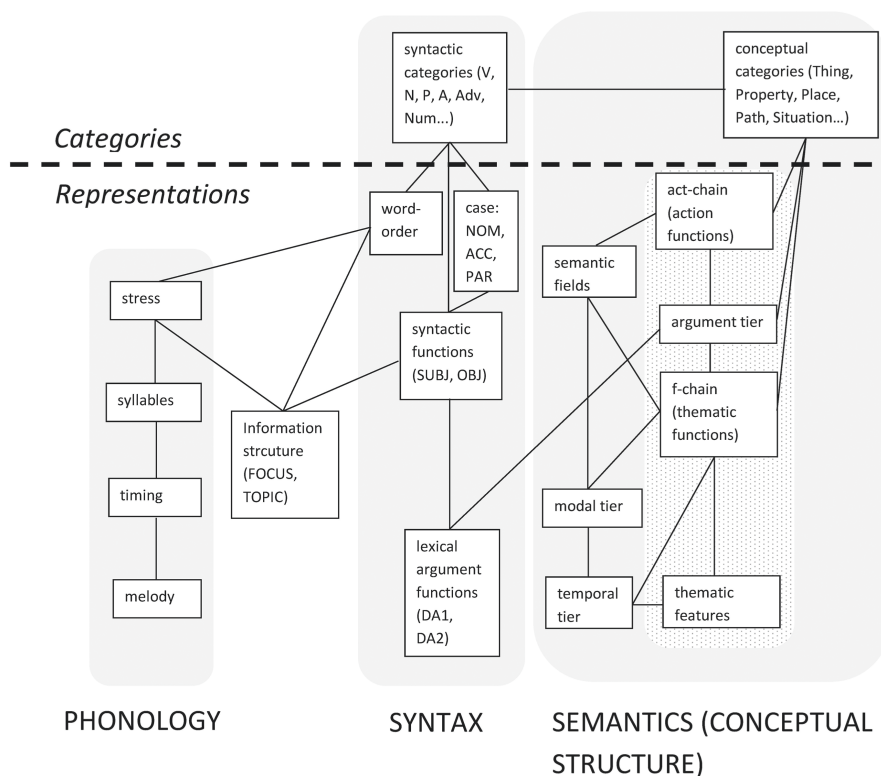
Figure 1: Micro-modular system of (the Finnish) language (Nikanne 2018: 38, Fig. 3-1).

For the sake of the argumentation presented in the chapter, it is not necessary to go to all details of the system. In order to make the overall picture more transparent the descriptions will be simplified as much as possible. However, the reader should to bear in mind the fact that the different representations of a linguistic expression form a whole. Therefore, the links between the representations must be recognized and allowed by the grammar of the language in question.

3.2 Levels of argument structure and the treatment of null subjects

The main idea is that the interpretation is ultimately based on the lexical conceptual structure of the predicate verb, and no empty syntactic element for the "missing" subject argument is needed. The interpretation procedure has three separate levels of argument structure (9).

(9) Levels of argument structure (the interpretation procedure)
  a. The GF-system: *grammatical functions* subject and object, based on the syntactic structure.
  b. The DA-system: *direct arguments* DA1 ('logical subject') and DA2 ('logical object'), based on the lexical entry of the predicate verb.
  c. *Thematic arguments* ('semantic roles'), based on the conceptual structure.

The DA-system (Nikanne 2018: 146; see also Nikanne 1995, 1997) can be defined as in (10).

(10) DA-system
  a. DA is a function that intermediates between the argument level and the grammatical functions; a DA
    (i) selects one argument among the arguments that are generated by the lexical f-chain or listed in the lexical entry;
    (ii) selects one grammatical function (subject or object).
  b. The primitives of the DA-system are of two kinds: DA1 and DA2. These primitives are hierarchically ordered: DA1 > DA2.

According to Nikanne (2018: 28), selection is a property of an operator (feature, function) to pick the element to its scope. The order of the arguments is based on the so-called *zones* that are the main conceptual structure formation rules (Nikanne 1990, 2018). Following this theory, the conceptual structure is ordered in three zones, see Table 1.

Table 1: Three zones of conceptual structure.

Zone number and description	Zone 3 Causative zone	Zone 2 Non-causative situation zone	Zone 1 Location zone
Conceptual category	Causative events and states	Non-causative events and states	Places and paths
Conceptual functions (as in Jackendoff 1983, 1990)	CS, INC	GO, BE, STAY, EXT, ORIENT	TO, TOWARD, FROM, AWAY-FROM, VIA, AT, IN, UNDER, etc.
Thematic role	Causer	Theme	Goal, Source, Route, Location

Here is a simple example that gives an idea of the zones (see (11)):

(11)

Zone		Zone 3		Zone 2		Zone 1	
Corresponding thematic role or conceptual function	Causer	Causative predicate (corresponding to the function CS)		Theme	non-causative situation predicate (corresponding to the function GO)	Path predicate (corresponding to the functions TO-IN)	Goal
Example	*John*	*made*		*me*	*go*	*into*	*my room*

In the example above, the predicates and arguments for each zone are expressed by separate words. This is of course not always the case. For instance, the verb *send* includes both causative and non-causative situation functions (CS-GO), e.g. *John sent me to my room* (see (12)).

(12)

Zone		Zone 3		Zone 2		Zone 1	
Corresponding thematic role or conceptual function	Causer	causative predicate (CS) and non-causative predicate (GO) combined.		Theme	Non-causative situation predicate (GO) included in the lexical conceptual structure of *send*.	Path predicate (corresponding to the functions TO-IN)	Goal
Example	*John*	*sent*		*me*		*into*	*my room*

In the verb *enter*, both the location zone functions TO-IN and the non-causative situation function GO are included in the verb, e.g. *John made me enter my room* (13).

(13)

Zone		Zone 3		Zone 2		Zone 1	
Corresponding thematic role or conceptual function	Causer	Causative predicate (corresponding to the function CS)		Theme	Non-causative situation predicate (GO) and location functions (TO-IN) combined	Path function (TO-IN) included in the lexical conceptual structure of *enter*.	Goal
Example	*John*	*made*		*me*	*enter*		*my room*

The lexical f-chain of a word is the fragment of a well-formed chain of conceptual functions stored in the lexical entry of that word. The functions select each other so that the causative functions select non-causative functions, which, in turn, may select path or location functions. Therefore, the functions select other always in the order indicated in (14) (f signifies "function" and ζ "zone").

(14) $f\zeta \rightarrow f\zeta/f\zeta{-}1$

Thus, a function may select another function from its own zone (1, 2, or 3) or a zone whose number is one smaller. However, zone 2 function can only select a function from zone 1. According to Nikanne (1990, 2018), the zone 2 function is also the head of the f-chain and it is always obligatory in a well-formed f-chain. The form of the f-chain is, thus (15), in which, * stands for none or one $f\zeta$ or several selections of the form $f\zeta \rightarrow f\zeta$.

(15) $f3^{\star} \rightarrow f2 \rightarrow f1^{\star}$

The f-chain functions select thematic arguments. Consequently, the hierarchy of the arguments is as in (16).

(16)   Causer > Theme > Location zone arguments

Selection means that linking of elements on separate levels of the system is asymmetrical: it has a direction. Selection in this case can be characterized as picking order: DA1 (logical subject) has the right to pick the most desirable thematic argument according to the hierarchy in (11), and DA2 (logical object) is the second in the picking order and has the right to pick the most desirable argument that is available.

The DA-functions select the grammatical functions SUBJ and OBJ. Thus, the conceptual structure functions (Causer, Theme, Location) are linked to the grammatical functions via DA-functions, not directly (Nikanne 1995, 1997). A similar type of intermediate system is suggested also by Culicover & Jackendoff (2005).

The GF-system is the system of grammatical functions, and it can be defined as follows (Nikanne 2018: 147).

The GF-system
a. GF is a function that selects an NP and is selected by a DA.
b. The primitives of the GF-system are SUBJ (subject) and OBJ (object).
c. SUBJ and OBJ are hierarchically organized: SUBJ > OBJ.

Thus, the DA-level functions select the grammatical functions in the same manner as they select the lexical conceptual structure arguments: DA1 is before DA2 in the picking order, and SUBJ is a more desired grammatical function than OBJ.

The linking between the GF-level and the DA-level follows the following principles (Nikanne 2018: 147). The linking has three steps, which are represented in (17).

(17) *Argument linking principles*
       Step 1
       a. If an argument is not implicit, it is a potential direct argument.
       Step 2
       b. The potential direct argument selected by the function with the scope over the largest number of other functions in the same lexical f-chain is linked to DA1.
       c. The potential direct argument selected by the function with the scope over the next largest number of functions in the same lexical f-chain is linked to DA2.
       d. If the selection of direct arguments by DA1 and/or DA2 is specified in the lexicon, the lexically specified selections overrule (b) and (c).
       Step 3
       e. Default linking between DAs and grammatical functions (in Finnish) is as follows: DA1 selects SUBJ and DA2 selects OBJ.

The direct arguments DA1 and DA2 are derived in the lexicon or, in exceptional cases, specified in the lexical entry. The derivation of the direct arguments from the lexical conceptual structure is given in Nikanne (2018: 135–175; see also Nikanne 1995, 1997).

It might seem that the intermediate lexical argument level is an unnecessary complication in the system, but I would like to point out that it actually makes the system simpler as a whole. In particular, the system of grammatical functions in Finnish is only based on two categories, i.e. SUBJ (syntactic subject) and OBJ (syntactic object). The syntactic arguments do not need to be marked in the lexicon. The assignment of the features (or linking of the intermediate subsystem of) logical subject and logical object to thematic structure arguments is based on a linking system; only exceptions of the system must be specified in the lexicon.

I also skip the analysis of the constituent structure of the sentence as it is irrelevant for the argumentation. Three things are necessary, i.e. the lexical argument structure, direct arguments and the syntactic arguments. The null-subjects of Finnish can be largely explained within the theory of linking between the direct arguments and syntactic arguments.

A simplified analysis of the sentence *Minä heitän sinulle pallon* [I(NOM) throw.1SG you(SG).ALL ball.ACC] 'I throw you a ball' with the 1st person singular subject *minä* 'I' being present is given in (18). The analysis is based on the lexicon. The direct arguments DA1 and DA2 follow from the form of the lexical entry. The direct arguments select the conceptual arguments on one hand and the grammatical functions on the other. If nothing else is specified, DA1 selects SUBJ and DA2 selects OBJ. The traditional terms 'logical subject' and 'logical object' express this idea that there is a default link between lexically based arguments and grammatical functions. There is nothing 'logical' in these arguments. The selection of conceptual structure arguments is based on a system of its own (see Nikanne 2018). (The arrow indicates selection: X → Y stands for 'X selects Y').[4]

(18) I   throw.1SG you.ALL   ball.ACC
       [NP *minä*]   [V *heitän*] [*sinulle*] [NP *pallon*] Syntax
       ↑   ↑
       SUBJ   OBJ   Grammatical functions (GF-level)
       ↑   ↑
       DA1   DA2   Lexical argument functions
                                                           (DA-level)
       ↓   ↓
       [I]   [BALL]   Lexical Conceptual Structure
                                                           arguments

3.3 Subject requirement in Finnish

In many languages, for instance in English, the syntactic subject is required by the grammar. In Finnish, this is not the case. This idea can be formulated as follows (19):

(19) *Subject Requirement Principle in Finnish*
      Syntactic argument SUBJ does not have to be linked to any lexical element
      if the null subject (i.e. lack of subject) is licensed by a principle of Finnish
      grammar.

In the *Subject Requirement Principle* given in (19), the term *grammar* refers to the Finnish language system in general, including the lexicon and the linking

---

4 For more on selection, see Nikanne (2018).

between the syntactic and the conceptual structures. The term *null subject* indicates that there is no overt or covert morpho-lexical element in the syntactic structure that is selected by SUBJ.

3.4 Comparison with a syntax-based theory of null subjects

It is worth pointing out that the conceptual semantics architecture of grammar presented above differs from minimalist theories as it does not assume that syntactic structures play a more central role than other parts of grammar.

According to Holmberg (2010: 89), "there are two ways to derive null subjects: one is by means of incorporation of a subject pronoun in T. In this case the null subject is a deleted copy in a chain headed by T. Definite null subjects can be derived in this way in consistent NSLs only. The other is by deletion of a pronoun in Spec, TP, subject to control from a higher clause."

The grammaticality of null subjects in the minimalist accounts is analyzed as feature checking which is further expressed as movements in the syntactic structure. In that way, it is possible to check that the surface syntactic structure is in accord with, for instance, the argument structure of the predicate verb. In the present account, a subject is not necessary in the syntactic theory if the system, as a whole, allows the interpretation (linking) of the syntactic structure to the conceptual structure. The fundamental thing in the interpretation is that the subject argument required by the lexical conceptual structure of the predicate verb can be found in the conceptual structure on the basis of the syntactic structure. If the system can link a syntactic structure with a "missing" counterpart of the subject argument it is fine. Thus, in those cases of null-subjects discussed in this chapter, there is no need to create a subject in the syntactic structure in the first place. A successful interpretation calls for (i) reading the form of the syntactic structure and recognition of the lexical items in it, (ii) the comparison of the syntactic structure to the lexical conceptual structures of the lexical items in it; and (iii) checking whether there are linking rules (regular linking and constructions) that can link the syntactic structure to a well-formed conceptual structure in the way that the lexical conceptual structures of the lexical items fit in it without collision.

Considering the fact that in the present theory there is no need to find a counterpart of the subject argument in the syntactic structure, there is no need for assuming any move (or chain) in the syntactic structure to link the subject to an appropriate syntactic argument position, either. Thus, even though the conceptual semantics account aims at the same thing as the minimalist accounts, the way they seem to follow radically different paths.

4 Analysis of null subjects in Finnish finite sentences

4.1 *Pro*-drop in 1ˢᵗ and 2ⁿᵈ person

The analysis of the *pro*-drop is simple, as given in (20). There is no selection link between the grammatical function SUBJ and syntax.

(20)
	did.1PL	today	stupid.thing.PL.PART
	[ᵥ *Teimme*]	[*tänään*]	[ₙₚ *tyhmyyksiä*]
			↑
SUBJ			OBJ
↑			↑
DA1			DA2
↓			↓
[WE]			[STUPID-THINGS]

The conceptual argument 'we' is understood because the finite verb is in the 2ⁿᵈ person plural form. Therefore, the DA1 can select the argument [WE] on the conceptual structure level. The grammatical function SUBJ does not get realized in the "overt" syntactic structure as there is no link. In other words, the subject of the sentence does not have any lexical counterpart in syntax. This analysis does not involve any abstract covert argument, like *pro*, either: nothing is nothing.

The *pro*-drop rule in Finnish can be stated as follows:

(21) **Licensing of null subject in 1ˢᵗ and 2ⁿᵈ person sentences**
If the predicate verb of the sentence is in the 1ˢᵗ or 2ⁿᵈ person and there is no syntactic subject in the sentence, the DA1 of the predicate verb is licensed if the argument selected by the DA1 has the same person and number features as the predicate verb.

The person inflection of the finite verb in Finnish reveals the content of the conceptual structure argument ('I,', 'we,' 'you.SG', 'you.PL'). This does not necessarily mean that if a language has a person agreement, it also automatically has a *pro*-drop. For instance, as Vainikka (1989) points out, in spoken Finnish *pro*-drop is much more limited than in Standard Finnish. In Standard Finnish, on the other hand, the 1ˢᵗ or 2ⁿᵈ person pronoun subject is used mostly for emphasizing the word. The person agreement, however, makes the overt subject redundant and opens the door for *pro*-drop.

In colloquial Finnish, however, in some constructions *pro*-drop is typical. For instance, in a context in which the speaker is starting a new topic in the conversation, as in (22).

(22) a. *Ostit   sitten   uuden   auton.*
       buy.2SG   then   new.ACC   car.ACC
       'So, you bought a new car.'
   b.   buy.2PL   then   [new car].ACC
             [V *Ostit*]   [Adv *sitten*]   [NP *uuden auton*]
                                    ↑
      SUBJ   OBJ
      ↑   ↑
      DA1   DA2
      ↓   ↓
      [YOU(SG)]   [NEW CAR]

This kind of *pro*-drop pattern may be based on a construction of its own but it, nevertheless, follows the general *pro*-drop principle.[5]

4.2 Passive sentences

In Finnish, passive sentences do not have a syntactic subject. The understood subject argument, logical subject is typically a person or several persons. The passive is used as in the examples in (23).

---

5 In principle, in Finnish imperative sentences could be analyzed in the same fashion as the 1st and 2nd person *pro*-drop sentences since in Finnish, the imperative mood also agrees with the subject in person and number. Only the 1SG form is missing in the standard Finnish paradigm of persons in the imperative mood. The imperative forms of *katsoa* 'look' (the stem of the verb is *katso-*) are presented below.

(i)   SG   PL
    1   -   *katsokaamme* look.IMP.1PL
    2   *katso*   look.IMP.1SG   *katsokaa*   look.IMP.2PL
    3   *katsokoon* look.IMP.3SG *katsokoot*   look.IMP.3 PL

However, languages that do not allow *pro*-drop still may have imperative sentences that do not have a syntactic subject. Therefore, I do not want to make any strong claims about the analysis of null subjects in the imperative sentences.

(23) a. *Nuorena   luetaan   kirjoja.*
     young.ESS   read.PTCP.PP book.PL.PART
     'People read books when they are young.'
   b. *Täällä   luetaan   kirjoja.*
     Here   read.PTCP.PP book.PL.PART
     'People read / are reading books here.' / 'Someone reads / is reading
     books here.'

In (23a), the interpretation is that the sentence expresses a general rule, and the
subject is generic 'one, anyone.' The sentence in (23b) may be used either as a
general rule ('in this place, people read books'), or an observation ('books are
being read here').

   The lexical argument structure of the verb *lukea* includes DA1 (logical subject)
and DA2 (logical object). The null-subject of a passive sentence is interpreted and
licensed according to the principle stated in (24), the relevant analysis presented
in (25).

(24) **Licensing of null subjects in Finnish passive sentences**
   If the predicate verb of the sentence is the passive form and there is no
   syntactic subject in the sentence, the DA1 of the predicate verb is licensed if
   the argument selected by the DA1 is ARB.

(25) a. young.ESS   read.PTCP.PASS.PST   book.PL.PART
        [*Nuorena*]   [$_V$ *luetaan*]   [$_N$ *kirjoja*]
                                                     ↑
                    SUBJ   OBJ
                    ↑   ↑
                    DA1   DA2
                    ↓   ↓
                    [ARB]   [BOOKS]
   b. Here   read.PTCP.PASS.PST   book.PL.PART
        [*Täällä*]   [$_V$ *luetaan*]   [$_N$ *kirjoja*]
                                                     ↑
                    SUBJ   OBJ
                    ↑   ↑
                    DA1   DA2
                    ↓   ↓
                    [ARB]   [BOOKS]

Just like in the *pro*-drop cases, the passive sentence has an understood subject argument. The passive morphology reveals that the understood subject argument is ARB.

4.3 Generic sentences

Even though there is no pro drop in the 3rd person in Finnish, there are subjectless finite sentences, as in (26).

(26) *Kirjoja   Lukee   mielellään.*
    book.PL.PART   read.3SG   with.pleasure
    'It is a pleasure to read books (for anyone).'

In the 3rd person, null subject in a finite sentence is allowed if the interpretation is generic '(any)one. The construction that licenses the lack of subject in the generic sentences can be formulated (informally) as follows:

(27) ***Licensing of null subjects in Finnish generic sentences***
    If the predicate verb of the sentence is in the 3rd person singular form and there is no syntactic subject in the sentence, the DA1 of the predicate verb is licensed if the argument selected by the DA1 is ARB.

The analysis of the sentence in (26) is as given in (28).

(28)   book.PL.PART   read.3SG   with.pleasure
           [$_N$ *kirjoja*]   [$_V$ *lukee*]   [$_{Adv}$ *mielellään*]
           ↑
    SUBJ OBJ
    ↑   ↑
    DA1   DA2
    ↓   ↓
    ARB   [BOOKS]

The pragmatic interpretations of generic sentences and the "intended" reference of the subject argument in different contexts is described in a thorough article by Laitinen (1995). The analysis above assumes that the conceptual interpretation of the generic sentence construction is uniform, and the content of the conceptual level argument is ARB, which can be translated as 'anyone, someone, people in general.' The pragmatic interpretations of ARB belong to language use.

In addition to the generic subject, there is a construction in Finnish that resembles the generic sentences. Namely, if the predicate verb is *sanoa* 'say' or *väittää* 'insist' (or some other verb) and it appears in the 3rd person plural form, it can be used in the same way as the English construction "They say that…".

(29) *Sanovat/Väittävät että olen   tullut   vanhaksi*.
     say.3PL/insist.3PL that be.1SG come.PTCP old.TRL
     'They say/insist that I have become old.'

The analysis of the sentence in (29) is given in (30).

(30)   say.*3*pl   that be.1SG come.PTCP old.TRL
           [$_V$ *sanovat*] [*että olen tullut vanhaksi*]
           ↑
     SUBJ OBJ
     ↑   ↑
     DA1 DA2
     ↓   ↓
     ARB [BOOKS]

This 3PL generic sentence is not as productive as the generic sentences in the 1st person singular. For the 3rd person singular, there are no restrictions of the predicate verb that can be used. In the 3rd person plural, on the other hand, the choice of the predicate verb is very limited. I assume that the null-subject in the 3rd person plural is licensed by a construction that specifies the lexical and morphological choices. Here, it is worth pointing out that the generic sentences in the 3rd person singular and the *pro*-drop cases in the 1st and 2nd person also specify the morphological form of the predicate verb. Thus, the borderline between constructions and regular grammar is not completely clear. According to Holmberg & Nikanne (1994, 2002, 2008), discussing the topic position of generic sentences and *pro*-drop sentences, there are differences, for instance, when it comes to the possible use of different expletive elements in the topic position (see also Nikanne 2018, 236–261.) However, those differences do not play a role in the interpretation of the null subject of these sentences.

4.4 Sentences with *weather*-predicates

In Finnish, *weather*-predicates may, but do not have to appear without a subject. A typical example of a Finnish *weather*-sentence with a null subject is (7), repeated

in (31). A simplified lexical entry of *sataa* 'rain' is provided in (32) (for a detailed analysis, see Nikanne 2018).

(31) *Sataa*.
       rain.3SG
       'It is raining.'

(32)   *sataa* (Theme: [WATER][I])

The superscript index I indicates that the Theme argument 'water' is implicit, and not linked to a DA nor, naturally, further to SUBJ or OBJ. As the only conceptual structure argument is lexically marked as implicit, the DA-level functions do not "see" it and it does not get selected by the DA-level functions. Consequently, it does not get linked to any grammatical function, either. The meaning of the sentence is understood even though the overt syntactic structure does not have any subject or object.

Even though the sentences with *weather*-predicates are somewhat difficult to analyze, the description given in (32) is very simple. There is no lexical argument (DA) in the lexical entry of *sataa* 'rain.' The understood Theme argument with *sataa* is 'water' if nothing else is specified. The linking between the Theme argument 'water' and the sentence is licensed by the lexical entry of *sataa*. The *weather*-predicate *sataa* 'rain' has an implicit argument 'water'. The Theme argument is the only argument of the verb *sataa*. I return to the analysis of these examples shortly.

(33) a. *Sataa   lunta*.
           rain.3SG   snow.PART
           'It is snowing. (Lit.: raining snow)'
      b. *Sataa   rakeita*.
           rain.3SG   hail.PL.PART
           'It is raining hail.'
      c. *Sataa   vettä*.
           rain.3SG   water.PART
           'It is raining.'

Even though *weather*-predicates do not need to have a subject in Finnish, they may have one, as shown in (34). The subject of *weather*-predicates may vary a

lot.[6] The following examples (picked from the Internet) show how it works. In the examples in (34), the subject 'snow' is in the nominative singular (*lumi*) and the predicate 'rain' is in the 3rd person singular (*sataa*). In the examples in (35), the subject is in the nominative plural (*lumet*), the predicate agrees with it and is in the 3rd person plural (*satavat*). If the subject is in the nominative plural, the sentence refers more to 'snow cover' than to the substance 'snow' (the relevant words *lumi* 'snow' and *sataa* 'rain' have been boldfaced for transparency).

(34) a. *Joskus*   **lumi**   **sataa**   *varhain, joskus   taas*
          sometimes   snow.NOM   rain.3SG   early   sometimes   again
          myöhään.
          late.
          'Sometimes snow falls early, sometimes late.'[7]

      b. *Erään   uskomuksen   Mukaan   pysyvä   **lumi***
          one.*gen*   belief.*gen*   according.to   permanent.*NOM*   snow.*NOM*
          sataa   *puolitoista   kuukautta   ensimmäisen*
          rain.*3SG*   one.and.half   months   first.GEN
          kurkiauran   näkemisen   jälkeen.
          wedge.of.cranes.GEN   seeing.GEN   after
          'According to one belief, the permanent snow will fall one and half months after one has seen the first wedge of cranes.'[8]

(35) a. *Uudet   **lumet**   **satavat**   ennen   kuin   vanhat*
          new.*PL.NOM*   snow.*PL.NOM*   rain.*3PL*   before   than   old.*PL.NOM*
          sulavat.
          melt.*3PL*
          'Then new snow will fall before the old snow will melt.'[9]
      b. *Miksi   **lumet**   **satavat**   juuri   Etelä-Suomeen?*
          Why   snow.*PL.NOM*   rain.*3PL*   just   South-Finland.*ILL*
          'Why is the snow falling just in Southern Finland?'[10]

---

6 See Kolehmainen (2010) for an excellent overview of weather expressions in Finnish.

7 https://www.ilmatieteenlaitos.fi/lumitilastot (accessed 27 April 2020).

8 https://www.savonsanomat.fi/ (15 October 2015; accessed 27 April 2020).

9 https://yle.fi/uutiset/3-5353003 (11 April 2012; accessed 27 April 2020).

10 https://www.hs.fi/kotimaa/art-2000002929334.html (9 November 2016; accessed 27 April 2020).

   c. ***Lumet   satavat*** *marraskuussa viikoksi   tai kahdeksi ja*
      snow.PL.NOM   rain.3PL   November.INE week.TRL or   two.TRL   and
      parhaat   kesäkelit   saapuvat   samanmittaiseksi
      best.PL.NOM   weather.PL.NOM   arrive.3PL   same.length.TRL
      jaksoksi   toukokuussa.
      period.TRL   May.INE
      'The snow falls in November for a week or two and the best summer
      weather arrives for a similar period in May.'[11]
   d. *8-vuotiaiden   hiihtokoulu alkaa, kunhan   **lumet***
      8-year.old.PL.GEN ski.school   starts   as.soon.as snow.PL.NOM
      satavat   *maahan.*
      rain.3PL   ground.ILL
      'The ski school for the 8-year old children will start as soon as the snow
      will fall on the ground.'[12]

Argument linking with *weather*-predicates is based on the Adjunct Principle
(see Nikanne 1990, 1995, 2018; and also Jackendoff 1990), which states that the
lexically specified implicitness of an implicit conceptual level argument may
be removed. Implicitness of an argument means that it is "invisible" for the
DA-selection and, consequently, cannot be expressed as an overt argument in
syntax. When the implicitness is removed, the argument is treated as any other
argument.

(36) [$_{NP}$ *lumi*] / [$_{NP}$ *lumet*]   [$_V$ *sataa*] / [$_V$ *satavat*]
      snow.NOM / snow.PL.NOM   rain.3SG / rain.3PL
      ↑
      SUBJ
      ↑
      DA1
      ↓
      [WATER]^I

According to Nikanne (2018: 165), the partitive NP *lunta* 'snow.PART' can be ana-
lyzed as the subject:

---

11 https://shl.fi/tag/luistelu/page/2/?sovellus=2 (Sydän-Hämeen Lehti, 27 March 2017; accessed
27April 2020).
12 https://www.facebook.com/pg/iisalmenvisanhiihtojaosto/posts (Iisalmi Visa Cross-Country
Ski Division, posted on 20 Oct. 2019 (accessed 27April 2020).

(37) [_V *satoi*]   [_{NP} l*unta*]
     rain.PST.3SG   snow.PART
                   ↑
                   SUBJ
                   ↑
                   DA1
                   ↓
                   [SNOW]

This is not a very radical analysis, as in Finnish the partitive is used for marking the subject of an existential sentence, as in (38). This is the standard analysis of these sentences (e.g., see Hakulinen et al. 2004).

(38) *Kadulla*   *kävelee*   *ihmisiä.*
     street.ADE walk.3SG people.PL.PART
     'There are people walking on the street.'

Another construction that should be mentioned at this point is the one involving the argument structure of the verb *sataa* 'rain'. Let us consider the examples that refer to the first snow (cover) of the year, as in (39).

(39) a. *Missä*   *vaiheessa* **satoi**   **lumen**   *ja*   *missä*
          what.INE point.INE rain.PST.3SG snow.ACC and what.INE
          vaiheessa tulivat ne   *Kovat* *pakkaset, joista*   *talvisodan*
          point.INE came those severe frosts   of.which winter.war
          aika *oli*   *niin kuuluisa?*[13]
          time was so   famous
          'At what point did it snow and at what point came the severe frosts for which the winter war time was so famous?'
    b. *Tänään* **satoi**   **lumen**   *ja* *peitti*   *kuoppien*
          today   rain.PST.3SG snow.ACC and cover.PST.3SG path.hole.PL.GEN
          näkyvyyden.
          visibility.ACC
          'Today, it snowed and the snow covered the visibility of the path holes.'[14]

---

13 Eeva Kilpi. 1989. Talvisodan aika: lapsuusmuistelma. ['The Winter War time: Childhood memories']. Helsinki: WSOY.
14 https://www.facebook.com/tieno956/posts/1876522625794438/ (accessed 28 April 2020).

    c. *Onpa   kiva, että* **satoi**   **lumen**.
        be.3sg nice   that rain.PST.3SG snow.ACC
        'It is so nice that it snowed.'[15]

The accusative Theme *lumen* 'snow.ACC' in (39) is the object and the sentences do not have a subject. This use of the accusative argument with *sataa* is exceptional and may be accounted for by assuming that this type of structure is licensed by a separate linking principle, consider (40).

(40)   ***Interpretation of the accusative form of lumi 'snow' with the verb sataa***
        If the predicate verb of the sentence is in the 3rd person singular form and there is no syntactic subject in the sentence, the DA1 of the predicate verb is licensed if the argument selected by the DA1 is ARB.

The DA1 of the verb *sataa* may select OBJ if the syntactic object is *lumen* 'snow. ACC', referring to the first snow of the year. Kolehmainen (2010) points out that the Source argument of *sataa* 'rain' may appear as the subject of the sentence (see 41).

(41)   *Pilvi   satoi   vettä.*
        cloud rain.PST.3SG water.PART
        'The cloud was raining water.'
        (Kolehmainen 2010: 10)

This example is licensed based on the assumption stated in (42)

(42)   ***Linking of syntactic subject to thematic role Source with verbs of 'flow' type***
        If the predicate verb has the meaning 'flow', e.g. *sataa* 'rain,' *valua* 'flow', *vuotaa* 'leak, spill', *tihkua* 'trickle', the thematic argument Source of the lexical conceptual structure of the verb may be selected by DA1 and the Theme argument by DA2.

This type of licensing applies to examples in (43c) and (44c). Examples (43b) and (44b) are roughly synonymous with those in (43c) and (44c), respectively. In turn,

---

15 https://rikkaruohoelamaa.blogspot.com/2015/11/ensilumi-satoi-yolla.html (accessed 28 April 2020).

(43a) and (44a) have a partitive subject and are analyzed as existential sentences (see 38).

(43) a. *Vettä   valuu.*
      water.PART   pour.3SG
      'Water drains.'
   b. *Silmistä   valui   vettä.*
      eye.3PL.ELA   pour.3SG   water.PART
      'Water was pouring from the eyes.'
   c. *Silmät   valuivat   vettä.*
      eye.NOM-PL   pour.3PL   water.PART
      'The eyes were pouring water.'

(44) a. *Verta   vuotaa.*
      blood.PART   leak.3SG
      'Blood is leaking.'
   b. *Haavasta   vuotaa   verta.*
      wound.ELA   leak.3SG   blood. PART
      'Blood is leaking from the wound.'
   c. *Haava   vuotaa   verta.*
      wound   leak.3SG   blood.PART
      'The wound is bleeding.'

5 Conclusion

In this chapter, I have discussed the nature of the concept of the term 'subject.' In the micro-modular theory (see Nikanne 2018), 'subject' can be divided into four different levels. These are listed below as (45).

(45) (a) *Overt subject*: an argument that is represented by a lexical item in the syntactic structure.
   (b) *Grammatical function*: SUBJ (GF-level).
   (c) *Lexically determined*: 'logical subject' (DA-level function, typically DA1).
   (d) *Conceptual level argument* that is linked to the overt subject via a DA-level function and the GF-level function SUBJ.

The analysis focused on finite root sentences in Finnish. Finnish does not have a requirement of overt syntactic subject and the "missing" syntactic subject is analyzed as non-existent (i.e., there is no covert syntactic element functioning as the syntactic subject). In order to be licensed, a syntactic structure must have a legal (system-based) interpretation to a well-formed conceptual structure and the lexical conceptual structures of the lexical items used in the syntactic structure must fit in that conceptual structure. Thus, if the syntactic structure with no syntactic subject has an interpretation, it is licensed. The aim of this chapter was to suggest principles and mechanisms that license the null subjects of Finnish finite root sentences.

Unlike many other languages, Finnish does not have a requirement of syntactic subject. Since the overt subject is not required, it is natural for the Finnish grammar to have several kinds of finite root sentences that lack an overt subject. Thus, *pro*-drop in the 1st/2nd person and the passive belong to the regular grammar of Finnish. The generic 3rd person singular sentences are productive, but their licensing is based on a construction that specifies the linking between the lexically determined subject argument and the content of the conceptual level argument. Additionally, the generic sentences in 3rd person plural require predicates that belong to a limited group of verbs of 'saying'. In turn, weather predicates (e.g. *sataa* 'rain') have a conceptual level argument that is marked implicit and, therefore, they need not have a subject. Finally, weather predicates may appear with subjects that are licensed by a variety of constructions.

References

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Hakulinen, Auli, María Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho. 2004: *Iso suomen kielioppi* [Great Finnish Dictionary]. SKST 950. Helsinki: Suomalaisen Kirjallisuuden Seura.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders & Urpo Nikanne. 1994. Expletives and subject positions in Finnish. In Javier Martin Gonzalez (ed.), *Proceedings of the North East Linguistics Society* 24, 173–187. Amherst, MA: University of Massachusetts.

Holmberg, Anders & Urpo Nikanne. 2002. Expletives, subjects, and topics in Finnish. In Peter Svenonius (ed.), *Subjects, Expletives, and the EPP*, 71–105. Oxford and New York: Oxford University Press.

Holmberg, Anders & Urpo Nikanne. 2008. Subject doubling in Finnish: The role of deficient pronouns. In Sjef Barbiers, Olaf Koeneman, & Marika Lekakou (eds.), *Microvariations in Syntactic Doubling*, 325–349. Bingley: Emerald.

Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1990. S*emantic Structures*. Cambridge, MA: MIT Press.

Kolehmainen, Leena. 2010. Sääverbien syntaksia ja semantiikkaa: semanttiset roolit, osallistujien vaihteleva käsitteistäminen ja sääverbien vaihteleva valenssi [Syntax and semantics of *weather*-verbs: semantic roles, varying conceptualization of participants, and varying valence of *weather*-verbs]. *Virittäjä* 114. 5–38.

Laitinen, Lea. 1995. Nollapersoona. [Null person]. *Virittäjä* 99. 337–358

Nikanne, Urpo. 1990. *Zones and Tiers: A Study of Thematic Structure*. Helsinki: Finnish Literature Society.

Nikanne, Urpo. 1995. Action tier formation and argument linking. *Studia Linguistica* 49. 1–31.

Nikanne, Urpo 1997. Lexical-conceptual Structure and Syntactic Arguments. Helsinki: SKY yearbook 1997.

Nikanne, Urpo 2005. Constructions in conceptual semantics. In Jan-Ola Östman & Miriam Fried (eds.), *Construction Grammars: Cognitive Grounding and Theoretical Extensions*, 191–242. Amsterdam & Philadelphia: John Benjamins.

Nikanne, Urpo 2018. *Conceptual Semantics: A Micro-modular Approach*. Amsterdam & Philadelphia: John Benjamins.

Paulsen, Geda. 2011. *Causation and Dominance: A Study of Finnish Causative Verbs Expressing Social Dominance*. Åbo: Åbo Akademi University Press.

Petrova, Oksana. 2011. *'Of Pearls and Pigs': A Conceptual-semantic Tiernet Approach to Formal Representation of Structure and Variation of Phraseological Units*. Åbo: Åbo Akademi University Press.

Pörn, Michaela. 2004. *Suomen tunnekausatiiviverbit ja niiden lausemaiset täydennykset* [Finnish emotive causative verbs and their sentential supplements]. Helsinki: Finnish Literature Society.

Vainikka. Anne. 1989. Deriving syntactic representations in Finnish. Amherst: University of Massachusetts PhD dissertation. https://finnishsyntax.files.wordpress.com/2015/11/vainikka1989.pdf (accessed on 9th October 2020).

Gréte Dalmi

9 Who on earth is *pro*? – Licensing null arguments in Hungarian matrix and dependent clauses

1 Introduction

This chapter[1] investigates the licensing conditions on null arguments (i.e. *pro*'s in the sense of Rizzi 1982, 1986) with the individual reference and the generic reference interpretations in Hungarian finite matrix and dependent clauses. Hungarian null arguments with the individual reference interpretation (referred to as *referential null arguments* in the generative syntactic literature) resemble their lexical pronominal counterparts in that they are free in matrix clauses, but they behave like bound variables in dependent clauses. In the latter case, they need to be bound (i.e. c-commanded and co-indexed) by an antecedent in an adjacent clause (see É.Kiss 1987, 2002; Kenesei 1989 for pronominal binding in Hungarian).

While it is sufficient for lexical pronominal arguments with the individual reference interpretation appearing in dependent clauses to be syntactically licensed, their null counterparts must also be semantically licensed by a discourse-related operator in the left periphery (i.e. the C-domain since Rizzi 1997), in the sense of Sigurðsson (2004) and Bianchi (2006).

The licensing mechanism proposed in this chapter is different from Frascarelli's (2007, 2018) Topic Criterion, according to which null pronominals must be licensed by the Aboutness-Shift Topic in all syntactic environments. Hungarian sentences have a rich left periphery with an Operator Field in it (see Bródy 1995). The locus of semantic licensing is the Operator Field.

Generic arguments (often called *impersonal* or *non-referential* in the literature) also have free and bound occurrences in Hungarian. The free occurrences are lexical, *az ember* 'one',[2] while the bound variable occurrences are null and they will be referred to here as pro_{GN}. In contradistinction to pro_{arb} (the null counterpart of 3PL generic exclusive *az emberek* 'people' (see Tóth 2009; Bródy 2013),

which is licensed by a quantifier in the Operator Field, generic inclusive pro_{GN} is a first person-oriented variable, requiring a coreferential antecedent in an adjacent clause. It must be semantically licensed by GN. GN is an intensional operator ranging over possible worlds (Moltmann 2006, 2010), and located in the Speech Act Participant Phrase (SAPP) within the C-domain (see Bianchi 2006; D'Alessandro & Alexiadou 2003; Dalmi 2013, 2014, 2017, 2018).[3] Generic lexical and null arguments must both be in the scope of GN.

In the typological scale established by Roberts & Holmberg (2010), Hungarian resembles radical NSLs in that in finite indicative clauses any argument can remain silent as long as it can be interpreted at a later stage of the derivation. At the same time, Hungarian also differs from radical NSLs in disallowing free variation between 3SG individual reference null arguments and 3SG generic reference null arguments. This places Hungarian somewhere between partial and radical NSLs in the typological scale of Null Subject Languages (NSLs) proposed by Roberts & Holmberg (2010:12).

1.1 Null arguments in generative grammar

Rizzi's (1982) seminal work on Italian null subjects reveals that languages show parametric variation in allowing subjects to remain silent. The parameter stating this cross-linguistic variation has become known as the Null Subject Parameter (see also Jaeggli & Safir 1989). It distinguishes languages in which subjects must be lexical from those in which they can also be null. The conditions under which an argument can remain silent (i.e. null) vary from one language to another. In some languages any argument with the individual reference interpretation can remain silent because the rich agreement morphology ensures its semantic reconstruction at LF (see Chomsky 1981, 1995; Rizzi 1982, 1986). In other languages, context is a sufficient criterion for "argument drop" (see Huang 1984, 1989). Depending on whether arguments are allowed to be null or not, human languages can roughly be divided into NSLs and non-NSLs (see Jaeggli & Safir 1989).

---

3 Ritter & Wiltschko (2019) outline a system of so-called impersonal pronouns which lack the Speech Act layer altogether. The present proposal relies on Moltmann's (2006, 2010) semantic account of first person-oriented generic *one*, expressing the speaker's attitudinal modality. First and second person pronominals are taken to be Speech Act Participants (SAPs), hence they are automatically assigned the [+SAP] semantic feature, which must be licensed in the SpeechAct-Participant Phrase (SAPP) projection (see Bianchi 2006) on the left periphery of the clause.

NSLs have recently been further divided into four major subclasses, according to the range of null subjects they allow (see Roberts & Holmberg 2010; Holmberg & Sheehan 2010 for details), with each class characterized by a range of properties. The subclasses are claimed to constitute a typological scale (see Roberts & Holmberg 2010: 12):

(1)   Expletive NSLs ⊃ Consistent NSLs ⊃ Partial NSLs ⊃ Radical NSLs

Languages on the right of the scale are in a proper subset relation with languages on their left. Since this classification, the theory of null subjects has undergone robust changes due to the accumulation of data from more and more NSLs (see Cognola & Casalicchio 2018). The intense study of the distribution, structure and licensing of generic inclusive and generic exclusive null subjects has led to a more fine-grained analysis of null subjects (see Krzek 2012 for Polish; Dalmi 2013, 2014, 2017 for Hungarian; Bizzarri 2015; Tsedryk 2013 for Russian; Holmberg & Phimsawat 2016 for Thai; Frascarelli 2018 for Italian; and Barbosa 2009, 2019 for European and Brazilian Portuguese, among others). These new developments have turned out to be crucial for the theory of null subjects and the typology of NSLs.

1.2 Genericity and clusivity[4] of null arguments

In the inventory of null arguments, some are assigned the individual reference interpretation, others receive the generic reference interpretation.[5] For example, the reference set of 2SG *pro* 'you' is a singleton, while 2SG generic *pro* (you$_{GN}$) refers to an unspecified group of entities including the speaker.[6] If the reference set of a generic expression includes the speaker, we talk about the generic inclusive interpretation. Generic arguments whose reference set excludes the speaker receive the generic exclusive interpretation (see Holmberg 2005, 2010).

---

4 Clusivity is a term used to distinguish inclusive and exclusive plural pronouns (see Simon 2005; Cysouw 2008).
5 Lexical and null arguments with the individual reference interpretation are often called *referential,* while lexical and null arguments with the generic reference interpretation are traditionally called *impersonal* (see Cinque 1988; Cabredo-Hofherr 2003; Egerland 2003 and the references therein). These terms do not capture the relevant interpretive differences between the two types of null arguments (see Holmberg 2005, 2010) and will not be used in the chapter.
6 Notice that reference to an unspecified group of entities does not equal to lack of reference.

Languages show a great diversity in the linguistic expression of generic reference. In non-NSLs, generic inclusive and generic exclusive arguments are both lexical:[7]

(2) a. *One   knows   what   one   should   wear   in the Opera.*
    b. *People   know   what   they   should   wear   in the Opera.*
       (modelled on Moltmann 2006: 276)

Interestingly, Italian, which is a paradigm case of consistent NSLs in the typological scale established by Roberts & Holmberg (2010), also requires a lexical pronominal, *si*, to express generic inclusive and generic exclusive subjects:

(3) *Da   qui   si   vede   le montagne.*
    from   here   SI   see   the mountain
    'One can see the mountain from here.'
    (Italian, Chierchia 1995: 114)

(4) *In Italia   si   beve   molto   vino.*
    in Italy   SI   drink   much   wine
    'In Italy people drink a lot of wine.'
    (Italian, Chierchia 1995: 108)

Although morphologically identical, the two kinds of *si* (*one* vs. *people*) differ in a number of properties (see Cinque 1988; Chierchia 1995).

Finnish is a core representative of partial NSLs, in which 3rd person individual reference subjects must be lexical, as in (5), however, 3rd person generic reference subjects must be null, (6):

(5) *Hän* / *$\ast$pro$_{GN}$ herä-ä   aikaisin   kesällä.*
    he / (one)   wake-3SG up   early   summer-INE
    'He wakes up early in summer.' (individual reference interpretation)

---

(6) *Kesällä*   **hän* / pro_{GN}   *herä-ä*   *aikaisin.*
    summer-INE   he / (one)   wake-3SG up   early
    'In summer one wakes up early.' (generic inclusive interpretation)
    (Finnish, adopted from Holmberg 2005, 2010)[8]

Based on these observations, Roberts & Holmberg (2010) divide NSLs into four major subclasses (expletive, consistent, partial and radical), depending on the kind of null subjects they allow. In Part 2, the place of Hungarian within this typology will be investigated.

2 The place of Hungarian among NSLs

2.1 The inventory of Hungarian null subjects

As Dalmi (2013, 2014, 2017, 2018) notes, Hungarian has all the characteristic properties of NSLs in the sense of Jaeggli & Safir (1989), Sigurdhsson & Egerland (2009) and Roberts & Holmberg (2010). It allows expletive null subjects in meteorological sentences, as shown in (7), individual reference null subjects, as in (8), generic inclusive and generic exclusive null subjects, (9)–(10), and, finally, any other null argument in finite clauses, as is shown in (11).

(7) *Már*   *sötéted-ett*   pro, *amikor haza-ér-tek*
    already   darken-PST.3PL   EXPL when   home-reach-PST.3PL
    a gyerekek.
    the children
    'It was already getting dark when the children got home.'

(8) *Vera*_i *fél-t*   *hogy*   pro_{i/j} *le-kés-i*   *a*   *vonat-ot.*
    Vera   fear-PST.3SG Th that   (s/he) PFX-miss-PRS.3SG the   train-ACC
    'Vera feared that s/he would miss the train.'

---

8 Finnish uses a special passive construction to trigger the generic exclusive interpretation:

(i) Rankassa   syödään   hyvin
    in France   eat.PASS.PRTC.3SG well
    'People in France eat well.'
    (Finnish, Holmberg 2010: 203)

(9) *A   középkor-ban   az emberek*_{arb} *fél-tek   a villámlás-tól.*
the Middle Ages-INE the people   fear-PRS.3PL   the lightening-ABL
'In the Middle Ages people feared lightning.'

(10) *Itt   nem beszél-nek   pro*_{arb}   *magyar-ul.*
here   not speak-PRS.3PL (they/people) Hungarian-FORM
'They do not speak Hungarian here.'

(11) *Ha   meg-érkez-el   pro*₁ *fel-hív-hat-od   pro*₁ *pro*₂.
when PFX-arrive-PRS.2SG (you) PFX-call-POT-2SG (you) (him)
'When you arrive, you can call him.'

Hungarian can be characterized by the following distribution of lexical and null pronominal arguments:

(12) *The distribution of lexical and null pronominal arguments in Hungarian*
  i.   Hungarian has expletive null subjects in meteorological sentences;
  ii. Hungarian allows individual reference null arguments in all persons;
  iii. Hungarian 3SG and 3PL generic reference arguments can be either lexical or null;
  iv. Hungarian 3SG null arguments do not alternate between the individual reference and generic reference interpretations.

As Dalmi (2013, 2014, 2017) points out, Hungarian differs from the four major classes of NSLs established by Roberts & Holmberg (2010) in three crucial respects, briefly discussed below. These properties place Hungarian somewhere between partial and radical NSLs.

 Property 1. 3SG individual reference null pronominals may alternate with their lexical counterparts but no free variation is possible between the individual vs. generic interpretations of 3SG null arguments.

3SG individual reference null pronominals are licit without any lexical antecedent in matrix clauses, though they require an antecedent when they appear in dependent clauses, (13)–(14). In such cases they often alternate with their lexical pronominal counterparts.

(13) *Péter*_j *nem gondolja,   hogy ő*_j / pro_j   *át-megy*
Peter   not think.PRS.3SG that   he/ (he)   through-go.PRS.3SG
a vizsgá-n.
the exam-SPR
'Peter_j does not think that (he_j) will pass the exam.'

(14) $Ő_j$ / pro$_j$ *nem gondolja,*   *hogy* pro$_j$ / $ő_j$ *át-megy*   *a*
      he / (he)   not think.PRS.3SG   that (he) / he   through-go.PRS.3SG   the
      vizsgá-n.
      exam-SPR
      'He$_j$ does not think that he$_j$ / (*pro$_j$*) will pass the exam.'

Unlike consistent NSLs, Hungarian has no overt 3SG generic inclusive lexical pronoun at all. 3SG generic null arguments must be syntactically licensed by a lexical antecedent and semantically licensed by being in the scope of GN.[9] The different licensing conditions exclude any variation between 3SG individual reference and generic reference null arguments.

(15) *Az ember*$_{GN}$ *nem gondolja*   *hogy*   *$^\star ő_j$* / pro$_j$ *át-megy*
      the man   not think.PRS.3SG   that   3SG / (3SG)   through-go.PRS.3SG
      a vizsgá-n.
      the exam-SPR
      '*One*$_{GN}$ does not think that *he$_i$ / pro$_i$ will pass the exam.'

Property 2. 3sg generic reference arguments, on their first occurrence, must be free lexical expressions; any further occurrences must be null and must be bound by a 3SG lexical antecedent.

As the examples in (16)–(17) indicate, the second occurrence of a 3SG generic inclusive argument must be null; it needs to be licensed by an appropriate antecedent in an adjacent clause. Referential pronominal arguments do not alternate with *pro*$_{GN}$, and they do not serve as antecedents for *pro*$_{GN}$, (18).

(16) *Az ember*$_{GN}$ *nem gondol-ja*   *hogy*   **az ember*$_{GN}$ / *pro*$_{GN}$
      the man   not think-PRS.3SG   that   the man / (the man)
      át-megy   *a vizsgá-n.*
      through-go.PRS.3SG   the exam-SPR
      '*One*$_{GN}$ does not think that *one*$_{GN}$ will pass the exam.'

─────────────

9 Krifka et al. (1995) take GN to be a kind of universal quantifier. The concept of GN introduced by Moltmann (2006, 2010) and adopted in this chapter diverges from this tradition. Here GN is a complex operator that introduces first person-oriented epistemic modality to the proposition. Thus, 3SG generic inclusive arguments are first person-oriented, context-dependent items, ranging over accessible worlds, while 3PL generic exclusive arguments are bound by a quantifier in space and time (see Moltmann 2006, 2010).

(17) *Az ember*_{GN}   *nem gondol-ja,*   *hogy*   **ő*_i / pro_{GN}   *át-megy*
      the man   not think-PRS.3SG   that   he / (one)   through-go.PRS.3SG
      a vizsgá-n.
      the exam-SPR
      '*One*_{GN} does not think that *he_i / *one*_{GN} will pass the exam.'

(18) **Ő*_i / pro_i   *nem gondol-ja,*   *hogy*   *az ember*_{GN}   *át-megy*
      he / (he)   not think-PRS.3SG   that   the man   through-go.PRS.3SG
      a vizsgá-n.
      the exam-SPR
      '*He_i / (He) does not think that one_{GN} passes the exam.'

Property 3. 3SG generic inclusive null pronominals show no WCO effect when they interact with quantifiers due to the fact that they are in the scope of GN, which always takes widest scope.

Existentially quantified DPs show quantifier interaction with other quantifiers; however, generic inclusive DPs do not interact with true quantifiers, due to the fact that they are in the scope of the GN operator (see Moltmann 2006, 2010). While existentially quantified DPs show Weak Cross-Over (WCO) effect, no WCO-effect is found in the case of generic inclusive DPs in similar contexts.[10] Thus, the quantified DP *valaki* 'someone' can only bind the null pronominal as long as it c-commands it in (19):

(19) V*alaki-t*_i   *át-enged*   [_{DP} *a(z) ő*_i / pro_i   *(saját)*
      someone-ACC   through-let.3SG   the he.NOM / (he).NOM   own
      professzor-a]   *a doktori   vizsgá-n.*
      professor-POSS.3SG   the doctoral   exam-SPR
      "His_i (own) / *pro*'s (own) professor will let someone_i pass (at) the doctoral exam."

If the quantified DP is not in a c-commanding position, quantifier-variable binding is not licit anymore:

---

10 WCO-effect is typically found in sentences like (i) and (ii) in English:
(i) Some people_i like their_i mother. 'There exist x's such that for every y, if y is x's mother, then y likes y.'
(ii) ??Their_i mother likes some people_i. 'For every y there exists an x such that y is x's mother and x likes y.'
These sentences are problematic for the Proper Binding Condition (May 1977), which requires that a quantifier should precede and c-command all instances of the variable it binds at LF. In Weak Cross-Over cases the Proper Binding Condition is fulfilled at LF, however the sentences are still ungrammatical.

(20)   *Valaki*_i   *(saját) professzor-a*   *át-enged-i*
       someone.NOM   (own) professor-POSS3SG   through-let-PRS.3SG
       őt*_{i /}pro*_i   *a doktori*   *vizsgá-n.*
       he.ACC / (he.ACC) the doctoral exam-SPR
       'Someone_i's (own) professor will let him_i /*pro*_i pass (at) the doctoral exam.'

Given that GN takes scope over the whole proposition, no similar asymmetry in grammaticality is found with generic inclusive DPs:

(21)   *Az ember-t*_{GN} *át-enged-i*   *a*   pro_{GN}   *(saját)*
       the man.ACC through-let-PRS.3SG (the man.NOM) own
       professzor-a   *a doktori*   *vizsgá-n.*
       professor-POSS.3SG the doctoral exam-SPR
       'One_{GN}'s (own) professor lets one_{GN} pass (at) the doctoral exam.'

(22)   [_{DP} *Az ember*_{GN} *(saját)*   *professzor-a]*   *átenged-i*
       the man_{GN}.NOM own   professor-POSS.3SG through_let-PRS.3SG
       *az ember-t*_{GN} / **őt*_i/ **pro*_i   *a doktori*   *vizsgá-n.*
       the man_{GN}-ACC he.ACC/(he.ACC) the doctoral exam-SPR
       'One's (own) professor lets one / *him_i / *pro*_i pass (at) the doctoral exam.'

This follows from the fact that GN is not a quantifier but an operator. The GN operator always takes widest scope as it has the whole proposition in its scope (see Moltmann 2006).

    Now that we have become familiar with the existing types of null arguments in Hungarian, let's have a closer look at the distribution and licensing conditions of 3SG individual reference lexical and null arguments in argument, adjunct and matrix clauses.

2.2 The distribution of 3SG individual reference lexical and null pronominals in Hungarian matrix and dependent clauses

3SG individual reference null pronominals are claimed to behave exactly like 3SG individual reference lexical pronouns. They are free in matrix clauses but they behave as bound variables in dependent clauses, hence they obey the rules of pronominal binding (see É.Kiss 1987, 2002; Kenesei 1989). Indeed, 3SG individual reference lexical and null pronominals often appear in free variation when their antecedent is a quantified DP:

(23)   *Mindenki-t$_i$*   *meglep-ett*   [$_{DP}$ *az,* [$_{CP}$ *hogy*   *ő$_i$ /*   pro$_i$ *megbukott*]].
      everyone-ACC   surprise-PST.3SG   it   that   he / (he)   flunked
      'It surprised everyone$_i$ that he$_i$ [had] flunked.'

(24)   [$_{DP}$ *Az,* [$_{CP}$ *hogy*   *ő$_i$ /* pro$_i$ *megbukott*]],   *mindenki-t$_i$*   *meglepett.*
         it   that   he / (he) flunked   everyone-ACC   surprise-PST.3SG
      'That he$_i$ had flunked surprised everyone$_i$.'
      (Hungarian, Kenesei 1989: 229–230)

Moreover, in (25), where the antecedent is a referential DP, the co-reference conditions for lexical and null pronominals are the same as in (23) and (24). The experiencer argument of the matrix *psych*-predicate serves as an antecedent for the lexical and the null pronominal in the dependent clause because they are in the proper c-command configuration.[11,12]

(25)   [$_{DP}$ *Az,* [$_{CP}$ *hogy ő$_i$ /* pro$_i$ *megbukott*   *a vizsgá-n*]],
         it   that   he / (he)   failed   the exam-SPR
      [$_{CP}$ *meglepte /* *bosszantotta*   *Péter-t$_i$*].
      surprise.PST.3SG / annoy.PST.3SG   Peter-ACC
      'That he$_i$ had failed the exam surprised/annoyed Peter$_i$.'

Obviously, a lexical antecedent in the dependent clause cannot c-command a lexical or null pronominal appearing in the matrix clause, hence the sentence is unacceptable, as is shown in (26). Co-variation in acceptability here suggests that the co-reference conditions are identical for lexical vs. null pronominals.[13]

---

11 Hungarian subordinate *that*-clauses are complements to an XP, whose head may be lexical or null (see Kenesei 1994; Szécsényi 2009).

12 *Psych*-predicates have a VP-shell in which either the Experiencer or the Theme is more prominent. In the causative *psych*-predicate class, the Theme argument is more prominent. Backward Binding is a syntactic operation often found with *psych*-predicates and *picture*-nouns. Here the antecedent follows but c-commands its bindee (see Pesetsky 1987).

(i)   [That she$_i$ would be late for the lecture] worried Sue$_i$ a lot.

(ii)   It worried Sue$_i$ a lot [that she$_i$ would be late for the lecture].

13 Notice that the configuration in (27) cannot be an instance of backward binding, as the lexical antecedent is not in a c-commanding position.

(26) *Meglepte / Bosszantotta* *$ő t_j$ / *pro$_j$ [$_{DP}$ *az,* [$_{CP}$ *hogy* *Péter$_j$* *megbukott*
surprised / annoyed   him / (him)   it   that   Peter   failed
a vizsgán]].
the exam-SPR
'It annoyed him$_{i/}$ / (him)$_i$ that Peter$_j$ had failed the exam.'

In contrast to the above scenario, we find a different picture concerning the distribution of lexical vs. null pronominal arguments in (27) and (28). In (27), the null pronominal argument is licit both in the matrix and in the subordinate clause, while the lexical pronominal argument is out under the intended coreferential interpretation; in (28) the opposite distribution is observed.

(27) *Péter-t$_j$*   *meglepte / bosszantotta az, hogy* *$ő_j$ / pro$_j$ *megbukott*
Peter-ACC surprised / annoyed   it   that   he/ (he)   failed
a vizsgá-n.
the exam-SPR
'It surprised/annoyed Peter$_i$ that *he$_i$ / (he$_i$) had failed the exam.'

Coreferential subjects in neutral sentences are normally dropped. The use of the lexical pronoun usually suggests contrast or emphasis. This is absent in (27), hence the overt pronoun makes the sentence ungrammatical there. In (28), the lexical antecedent is in the extraposed subordinate clause, where it does not c-command the null pronominal and even LF-reconstruction proves insufficient. The structure can only be rescued by using the overt pronominal:

(28) [$_{DP}$ *Az,* [$_{CP}$ *hogy* *Péter$_i$* *megbukott* *a vizsgá-n]],*   *bosszantotta*
it   that   Peter   failed   the exam-SPR   annoyed
$ő t_i$ /*pro$_i$.
him / (him)
'That Peter$_i$ had failed the exam annoyed him$_i$ / (*him$_i$).'

Even this short digression can illustrate that the licensing conditions of lexical and null pronominal arguments in Hungarian dependent clauses differ substantially; their distribution is strongly conditioned by discourse-semantic factors.

2.3 The feature composition of individual vs. generic reference null arguments

The widely held view about null pronominal arguments cross-linguistically is that they are underspecified for *phi*-features; this is how they are distinguished from fully specified lexical pronominal arguments (see Cinque 1988; Cabredo-Hofherr 2003; Holmberg 2005, 2010; Holmberg & Sheehan 2010; Holmberg & Phimsavat 2015; Barbosa 2009, 2019; Ritter & Wiltschko 2019). In Finnish, a partial NSL, only 3SG generic pronominals can remain silent; 3SG individual reference subjects must be lexically expressed in matrix clauses; they can remain silent in subordinate clauses only under some well-defined conditions (see Roberts & Holmberg 2010: 6). This interpretive split between the overt, lexical 3SG pronominal subject and the silent 3SG pronominal subject has turned out to be a crucial distinguishing property of partial NSLs.

2.3.1 Holmberg's (2005, 2010) model

In his recent work, Holmberg (2005, 2010) explains the differences between the individual vs. generic reference interpretations of 3SG null pronominals by the presence or absence of the [+D] feature on the D_0 head. The [+D] feature ensures that the null pronominal receives the individual reference interpretation.[14] Finnish 1st and 2nd person null pronominals are full DPs in his account, which can be freely dropped:

(29) a. pro *puhu-n* *englanti-a.*
        (I) speak-1SG English-PART
        'I speak English.'
    b. pro *puhu-t* *englanti-a.*
        (you) speak-2SG English-PART
        'You speak English.'
        (Finnish, adopted from Holmberg 2005: 539)

In these sentences, person/number agreement morphology on the verb guarantees the referential interpretation of the null subject. 3SG null pronominals with the generic reference interpretation, on the other hand, are D-less φPs. They instantiate weak pronouns in the sense of Cardinaletti & Roberts (2002). The only

---

14 The EPP feature has been split into [D] and φ-features since Chomsky (1995).

true instantiation of *pro*, Holmberg (2005) claims, is found in AGR-less languages like Chinese, and here it is a bare nominal, a phi-less noun, $[_N pro]$. It is not clear what consequences such variation in the projectional complexity of null arguments has for the theory of grammar in general.[15] However, it surely cannot be maintained in Hungarian.[16]

2.3.2 Hungarian personal pronouns are full DPs

Hungarian lexical and null personal pronouns are structurally deficient DPs, with a D_0 head and without a specifier of DP (see Abney 1986 for English and Longobardi 1994 for Italian):

(30) $[_{DP} [_{Do} ő_i / pro_i [_{No} t_i]]]$ 'he / (he)'

In the present account, personal pronouns are generated under N_0 and perform N_0-to-D_0 movement, in the sense of Longobardi (1994). They refer to entities, and project as DPs. Their structural deficiency does not level them with demonstrative and possessive pronouns, which are premodifiers of nominal heads, and as such, are located in the pre-head domain of the DP. The syntactic and semantic features of lexical and null pronouns are shared between N_0 and D_0. An obvious example of this is the [±DEF] object agreement feature. As is typical in Ugric languages,[17] Hungarian verbs are morphologically marked for the definiteness of their direct object;[18] indefinite objects do not require an overt (in)definiteness agreement marker on the verb. Accordingly, lexical and null pronominal objects, just like full lexical DPs, store a [±DEF] feature on their N_0 head, which is licensed under the

---

15 See Déchaine & Wiltschko (2002) for a solution based on variation in the projectional complexity of Halkomelem lexical pronouns.

16 Nominal predicates have no DP projection, hence no [+D] feature; they are non-referential NPs (see Partee 1988; Rapoport 1995). This is not true for lexical or null pronominal arguments, which do have reference, whether it be individual or generic. Dvořak (2015) shows that *pro* is specified for gender in Czech, and Dalmi (2017) argues for a full DP-analysis of *pro* on the basis of [±DEF] lexical and null objects in Hungarian.

17 See Nikolaeva (1999) and Dalrymple & Nikolaeva (2011) on object agreement in Khanty.

18 Hungarian verbs are marked for the definiteness or indefiniteness of their object. This is shown by the [±DEF] feature in the glosses here. On Hungarian definite object agreement see Farkas (1987); Bartos (1997, 1999), (2000); Bárány (2015); Dalmi (1998, 2002); Coppock (2013); Coppock & Wechsler (2012).

D_0 head in the course of the derivation.[19] This excludes D-less null pronominals in this language. Within the verbal projection, the locus for licensing the [±DEF] feature of the direct object is ASPP, where the other object features of the verbal head are also licensed (see Dalmi 1998 and 2002).[20]

2.3.3 Generic inclusive and generic exclusive null arguments

Both 3SG inclusive and 3PL exclusive generic reference null arguments are co-indexed with their lexical antecedent *az ember* 'the man', and *az emberek* 'the people', respectively. They inherit the syntactic and semantic feature bundle of their respective lexical antecedent, hence they must be full DPs:

(31) *Az ember*_{GN} *divatosan* *öltöz-ik* *ha* *Párizs-ban* *él*
       the man_{GN}   fashionably dress-PRS.3SG if Paris-INE live.PRS.3SG
       pro_{GN}.
       (the man_{GN})
       'One_{GN} dresses fashionably if one_{GN} lives in Paris.'

---

19 On the syntactic category and feature composition of personal pronouns in Hungarian see Dalmi (1998, 2002). On a feature-geometric approach to pronouns in general see Harley & Ritter (2002).
20 As there is no distinct AGRoP projection in the functional domain of the verb (see Chomsky 1995), the only potential locus for checking the [+D] and [+DEF] features of objects within the functional layer is ASPP. Dalmi (1998, 2002) explicitly segregates intransitive verbs, which have no projection at all for licensing object agreement, from transitive verbs selecting an indefinite object, which do. In the alternative accounts, where such segregation is not made, transitive indefinite verb forms like *néz* 'watch' and intransitive verbs like *úszik* 'swim' can both have an indefinite object agreement feature on the verb, to be licensed in ASPP. Thus, there is nothing to exclude ungrammatical sentences like (ii):

(i) *Péter* *néz*   *egy film-et.*
     Peter watch.PRS.3SG.[-DEF] a film-ACC
     'Peter is watching a film.'

(ii) *Péter* *úszik*   (*egy folyó-t*).
     Peter swim.PRS.3SG.[-DEF] a river-ACC
     'Peter is swimming (*a river).'

Another problem that these alternative accounts cannot handle is verbs showing transitivity alternation. Thus, verbs like *eat, see, read, write* will receive the same treatment in their transitive indefinite uses as in their intransitive uses, unless they are taken to be distinct lexical entries.

(32) *Az emberek*_{arb} *divatosan* *öltöz-nek,* *ha* *Párizs-ban* *él-nek*
the people   fashionably dress-PRS.3PL if Paris-INE live-PRS.3PL
*pro*_{arb}.
(the people)
'People dress fashionably if they live in Paris.'

As is argued by D'Alessandro & Alexiadou (2003) for Italian, generic inclusive null arguments carry a [+GN] feature on their D_0 head. The [+GN] feature has to be licensed by the GN operator sitting in Speech Act Participant Phrase (SAPP) in the left periphery of the sentences (see Bianchi 2006). This automatically excludes an analysis of *pro* as an underspecified, minimal N_0 head (see Barbosa 2009, 2019 for this proposal in Brazilian Portuguese).

The 3PL generic exclusive DP *az emberek* 'people' and its null counterpart are quantificational (i.e. they are licensed in QP within the C-domain) and will not be discussed here in detail (see Tóth 2011 and Bródy 2013 for further details).

2.3.4 What is licensed where?

As opposed to consistent and partial NSLs, Hungarian allows 3SG generic inclusive lexical and null DPs to appear in any argument position in the clause; however, only [+human] null pronominal arguments can be interpreted as generic. The [+human] feature is stored on the N_0 head together with [+person], [+number] and [+case].

Ackema & Neeleman (2018: 130) attribute the [+human] requirement to the speaker-orientation of generic inclusive pronominal DPs. By the uniformity requirement on sets, the reference set of generic pronominal arguments must be semantically homogeneous. As generic inclusive lexical and null arguments are 1st person-oriented, i.e. they include the speaker (see Moltmann 2006), they cannot refer to [−human] entities, (33):

(33) *One is not allowed on board if one is overweight.*
(example from Ackema & Neeleman 2018: 130)

Likewise, the oddity of (34) stems from the fact that *one* can only refer to [+human] entities:

(34) *??If one is a Martian, one is not susceptible to human disease.*
(example from Moltmann 2006: 259)

Thus, we get the following feature bundles for individual reference and generic reference null DPs in Hungarian:

(35) a. Hungarian generic reference null pronominals

```
                DP
              /    \
           D_0      NP
            |        |
            |       N_0
            |        |
          pro_GN     t
          [±D]
          [±DEF]   [+person]
          [±GN]    [+number]
                   [+human]
                   [+case]
```
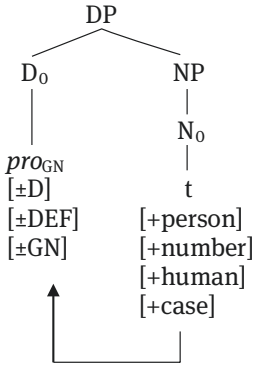
As we can see in (35a), the D_0 head hosts binary features. Null pronouns with the individual reference interpretation are [-GN], while generic null pronouns are [+GN]. Both types refer to entities, hence they also have a [+D] feature.

  b. Hungarian individual reference null pronominals

```
                DP
              /    \
           D_0      NP
            |        |
            |       N_0
            |        |
          pro_i      t_i
          [±D]
          [±DEF]   [+person]
                   [+number]
                   [+human]
                   [+case]
```
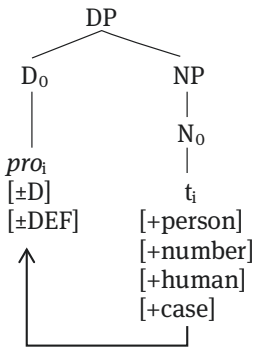
It is clear from the structural representation of Hungarian generic vs. individual reference null arguments given in (35a-b) that they do not differ to the extent that would explain their different licensing conditions. Indeed, in the present proposal this difference is not derived from the feature composition of generic vs. individual reference pronominal arguments alone; their licensing conditions will be discussed in Part 3.

3 Licensers and licensing conditions for Hungarian lexical and null pronominal arguments

As was shown in the preceding discussion, Hungarian displays both individual reference and generic reference null arguments. Their licensing conditions, however, differ substantially. In dependent clauses, individual reference *lexical pronominals* become licit merely via syntactic licensing; *null pronominals*, in addition, need to be also semantically licensed, i.e. they need to be in the scope of an operator.[21] I will remain agnostic as to whether this operator is an Iota-type operator in Partee's (1998) sense or something else. Suffice to say that the locus of licensing is the Operator Field within the C-domain. Now let us have a closer look at 1st and 2nd person pronominal arguments and their individual vs. generic interpretations.

3.1 Licensing 1st and 2nd person individual and generic reference null arguments

The syntactic licensing of 1st and 2nd person null pronominal subjects with the individual reference interpretation takes place in the TP functional projection in the standard way. The verbal head, which stores the [+person], [+number], [+case] features, in addition to the [+D] referentiality feature, enters into Agree relation with the T_0 head in the course of the derivation.

(36) a. *Lát-**om***   pro   *a fiú-t.*
           see-1SG.[+DEF]   (I)   the boy-ACC
           'I can see the boy.'
     b. *Lát-**od***   pro   *a fiú-t?*
           see-2SG.[+DEF]   (you)   the boy-ACC
           'Can you see the boy?'

1st and 2nd person null pronominal objects in Hungarian are specified for the [-DEF] feature, which is morphologically reflected by the absence of the definite object agreement suffix on the verb, as is shown in (37)–(39) below. The a. examples

---

21 This dual licensing procedure is similar to the one proposed for NPIs by Ladusaw (1996).

appear with the [-DEF] form of the verb, while the b. examples demonstrate that 1st and 2nd person pronominal objects are incompatible with the [+DEF] object suffix on the verb.[22] The example in (37) shows a 1st person null object, while in (38)-(39) a 2nd person null object is found.

(37) a. *Most lát*   pro$_1$ pro$_2$.
     now see-3SG[-DEF] (s/he) (I.ACC)
     'He can see me now.'

    b. *Most lát-ja   pro$_1$ pro$_2$.
     Now see-3SG[+DEF] (s/he.NOM) (I.ACC)
     'He can see me now.'

(38) a. *Most lát-tok*   pro$_1$ pro$_{2.}$?
     now see-2PL[-DEF] (you.NOM) (I.ACC)
     'Can you(PL) see me now?'

    b. *Most lát-játok   pro$_1$ pro$_2$?
     now see-2PL[+DEF] (you.PL.NOM) (I.ACC)
     'Can you(PL) see me now?'

(39) a. *Lát-unk*   pro$_1$ pro$_2$.
     see-1PL**[-DEF]** (we.NOM) (you.ACC)
     'We do see you.'

    b. *Lát-juk   pro$_1$ pro$_2$.
     see-1PL**[+DEF]** (we.NOM) (you.ACC)
     'We do see you.'

The indefinite object agreement morphology of the verb clearly identifies 1st and 2nd person objects as indefinite. This identification process requires a full DP.

---

22 The [±DEF] feature must be licensed by the verbal head in ASPP, together with the other object features of the verb.

(40)   *Indefinite null object agreement*

```
                    FinP
                 ┌────┴─────┐
              Fin          TP
               │       ┌────┴────┐
               │       T        AspP
               │     [+D]    ┌────┴────┐
               │    [+NOM]  Asp        vP
               │    [3SG]  [-DEF]  ┌────┴────┐
               │          [+ACC]  DP        v'
               │          [2SG]    │     ┌───┴────┐
               │                   │     v       VP
               │                   │         ┌───┴───┐
               │                   │         V      DP
               │                   │         │       │
             lát_i               pro_1      t_i     pro_2
          see.3SG.INDF          (3SG.NOM)          (2SG.ACC)
          'He can see you.'
```
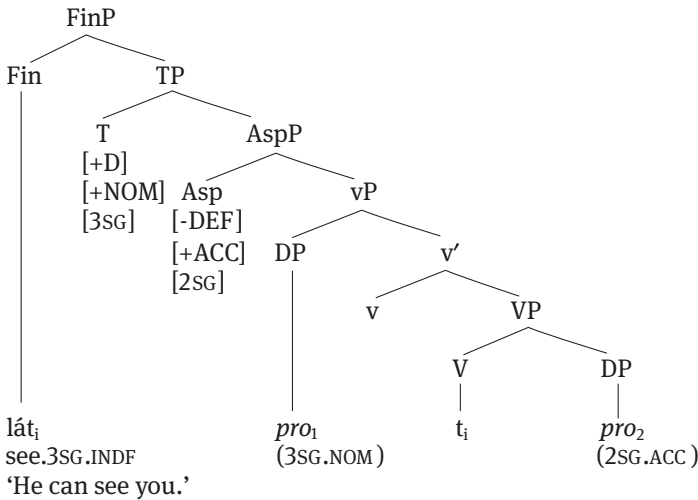
When the verbal head climbs up the tree in (40), it will have its syntactic features licensed by the T_0 and ASP_0 heads via Agree (see Dalmi 1994, 1998, 2002 for details), whereas its discourse-semantic features will be licensed by a discourse-semantic operator in the C-domain, in the sense of Sigurðsson (2004).

As opposed to Italian (see D'Alessandro & Alexiadou 2003; D'Alessandro 2007) or Icelandic (see Sigurðsson & Egerland 2009), Hungarian does not allow any alternation between the individual vs. generic reference interpretations of null pronominals. The feature that distinguishes generic reference null pronominals from individual reference null pronominals is [+GN]. Accordingly, 1PL and 2SG null pronouns can receive the generic inclusive interpretation on two conditions: (i) they enter the computation equipped with the [+GN] feature, which is licensed syntactically via Agree; (ii) they are licensed semantically by being in the scope of the GN operator. The locus of this operator is the SAPP functional projection within the C-domain (see Bianchi 2006):

(41)   *The C-domain of Hungarian clauses (Dalmi 2013, 2014, 2017)*

```
                    CP
            ┌───────┴───────┐
          Spec             C′
                     ┌──────┴──────┐
                     C            SAPP
                     │      ┌──────┴──────┐
                     │    Spec          SAP′
                     │             ┌──────┴──────┐
                     │            SAP          TOPP*
                     │             │      ┌──────┴──────┐
                     │            GN    Spec          TOP′
                     │                           ┌──────┴──────┐
                     │                          TOP           QP
                     │                                  ┌──────┴──────┐
                     │                                 Spec          Q′
                     │                                          ┌─────┴─────┐
                     │                                          Q         FOCP
                     │                                              ┌──────┴──────┐
                     │                                            Spec          FOC′
                     │                                              /\      ┌─────┴─────┐
                     │                                             /  \    FOC       FinP...
                    ha                                    elég sokáig élünk pro GN
                    if                                    enough long live (1PL)
              'If we GN live long enough...'
```
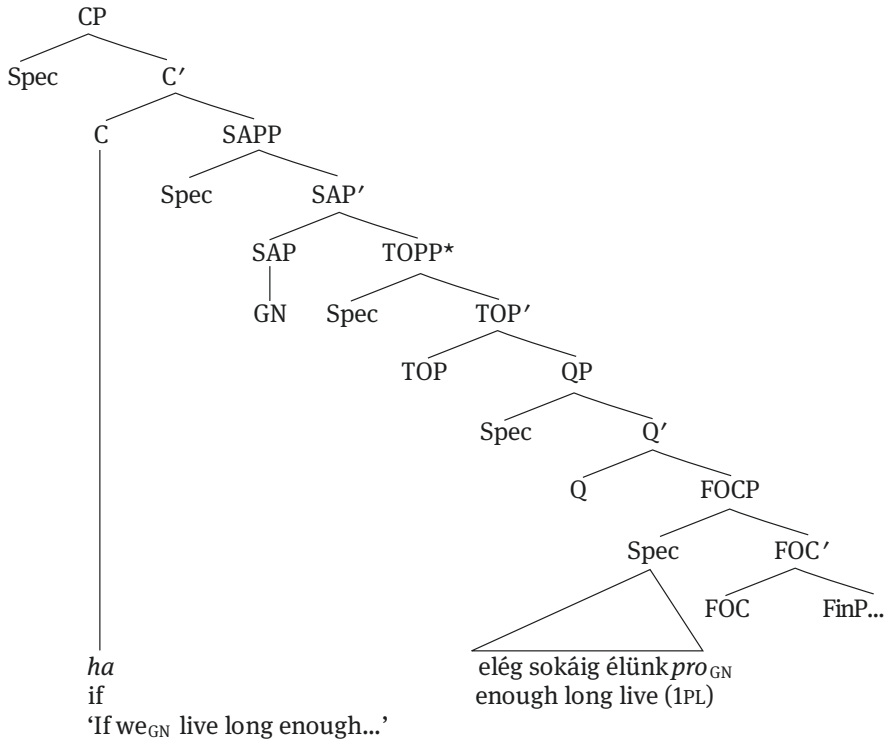
Thus, the generic reference interpretation of 1PL and 2SG person null arguments is conditional on syntactic licensing via Agree, as well as on semantic licensing by the GN operator in SAPP.

Given this sentence structure, the syntactic and semantic licensing of null arguments takes place in all persons without any reference to [+aboutness] topic in this language. In order to understand why this is so, it is inevitable to make a short digression to a recent theory put forward by Frascarelli (2007, 2018), concerning the licensing null subjects in Italian.

3.2 Frascarelli (2007, 20018)

Frascarelli (2007, 2018) claims that all null subjects need to be licensed by the so-called Aboutness-Topic in the C-domain of Italian sentences. She introduces the *Topic Criterion*, given here as (42):

(42) *Topic Criterion* (Frascarelli 2007)
a. [+aboutness] is connected with an EPP feature in the high Topic field
that yields a specific discourse-related property, namely "Aboutness".
b. The [+aboutness] Topic matches with an argument in the main clause
through Agree.
c. When continuous, the [+aboutness] Topic can be null (i.e. silent).

This licensing mechanism cannot be employed in Hungarian for several reasons.
First, in Hungarian, any argument of the verb can become null. Thus, it is not
surprising that we find dative experiencer null arguments:[23]

(43) *Az építkezés-en*   *kötelező volt*   *a védősisak   viselés*
      the construction site-SPR   obligatory was   the shield   wearing
      mindenki-nek.
      everyone-DAT
      'On the construction site, wearing a shield was obligatory for everyone.'

(44) *Az építkezés-en*   *kötelező volt*   pro$_{arb}$ *a védősisak viselés.*
      the construction site-SPR   obligatory was   (people.DAT)   the shield wearing
      'On the construction site wearing a shield was obligatory (for people).'

These arguments target a higher position than canonical subjects (see Cardina-
letti 1997 and 2004). Nonetheless, they are arguably not licensed by Topic but
by some modal operator in the C-domain (see Cardinaletti 1997, 2004 for Italian;
Benedicto 1995 for Russian; Dalmi 2000, 2005 for Hungarian).   Second, as
Frascarelli (2007) admits, her Topic Criterion, given in (42), holds for 3rd person
individual reference null subjects but does not extend to 1st and 2nd person null
subjects. Nor does it offer an explanation for the 3SG lexical vs. null pronominal
split in Finnish (Holmberg 2005, 2010). Third, TopicP in Hungarian is an optional
projection, which hosts [+specific] arguments (see É.Kiss 1987, 2002). TopicP is
outside the Operator Field. It differs from Focus and other operators in that it
has no semantic contribution of its own. It does have a pragmatic contribution,
such as modal reservation of the speaker or aboutness (see Szabolcsi 1981, 1994).
Finally, in Hungarian, the T_0 and ASP_0 heads syntactically license lexical pro-
nominal arguments with the individual interpretation, and operators like FOC, Q,
NEG can license null pronominal arguments semantically, in addition (see É.Kiss

---

23 On the VP-shell assumed for modal and psychological predicates with a dative experiencer
and a nominative theme see Benedicto (1995) for Russian, Jonsson (1997) for Icelandic and Dalmi
(2000, 2005) for Hungarian.

2002; Surányi 2004, 2006; and the references cited therein). Generic lexical and null arguments are invariably licensed by GN (see Moltmann 2006).

All these factors make the Topic Criterion immaterial in this language.

3.3 The present proposal

Starting out with a cartographic model of Hungarian sentence structure (see Dalmi 2005, 2010, 2014, 2017), the following *syntactic and semantic licensers* can be assumed for lexical and null pronominals in this language:[24]

(i) The GN generic operator semantically licenses both the generic inclusive lexical DP *az ember* 'one' and its null counterpart, pro_{GN}, by having them in its scope. 1PL and 2SG generic lexical and null pronominals are also licensed by GN in SAPP;

(ii) Quantifiers like MOST semantically license the generic exclusive lexical DP *az emberek* 'people' and its null counterpart, pro_{arb}, which are also licensed syntactically by the relevant functional categories in the T-domain *via* canonical syntactic binding;

(iii) The Focus operator semantically licenses lexical pronominals with the individual interpretation when they are contrasted or emphasized;

(iv) T(ense) and Asp(ect) invariably license lexical and null pronominals with the individual interpretation syntactically, via Agree.

(45) $[_{CP}...[_{SAPP}\textbf{GN}..[_{ForceP}..[_{TOPP}..[_{QP}..\textbf{Q}..[_{FOCP}..\textbf{FOC}..[_{FinP}...[_{TP}..\textbf{T}..[_{ASPP}\textbf{ASP}..[_{VP}..]]]]]]]]]]$
C-domain {--OPERATOR FIELD--} T-domain V-domain

In Hungarian, where any argument of the verb can remain silent, it is the verbal head that enters into cyclic Agree relation with various functional heads (á lá Bèjar & Rezac 2009) in the course of the derivation.[25] The T_0 head licenses the [+D], [+person], [+number], [+NOM] features of the verbal head. The [±DEF] defi-

---

24 TOPP in Hungarian is an iterative functional projection. The TOP_0 head has no quantificational force, it merely changes the truth conditions of the sentence but not its truth value (see Szabolcsi 1981, 1994). The T_0 head can license the relevant syntactic features of lexical and null pronominals.

25 Cyclic Agree is a syntactic process proposed by Bèjar & Rezac (2009) to handle the direct and inverse person/number agreement patterns of transitive verbs in Basque and other ergative/absolutive languages. The relevant features of the inner argument must be probed in the inner cycle. An important digression from the concept of Agree found in minimalist theory (Chomsky 1995) is that the cycles need not be *phi*-complete. Partial match produces a special feeding/bleeding effect of feature agreement in ergative languages.

nite object feature of the verbal head is licensed in ASPP along with the other object features (see Dalmi 1998, 2002, 2005 for this proposal). In addition to the usual ways of *syntactic* licensing, null arguments are also *semantically* licensed, by being in the scope of some operator. [26]

(46) *Licensing arguments in Hungarian*
    i.   Individual reference lexical pronominal arguments must be licensed syntactically, via Cyclic Agree between the verbal head and the relevant functional head, in the sense of Bèjar & Rezac (2009).
    ii.   Individual reference null pronominal arguments must be licensed both syntactically (via Cyclic Agree) and semantically (by the relevant operator within the Operator Field of the C-domain).
    iii.   Generic reference lexical and null arguments must be licensed semantically by the GN operator in SAPP, and syntactically, via Cyclic Agree between the verbal head and the relevant functional head, in the sense of Bèjar & Rezac (2009).

Temporal, aspectual, modal, negative and focus operators as well as quantifiers are all located in the Operator Field within the C-domain of Hungarian clauses (see Bródy 1990, 1995; É.Kiss 2002) and can license lexical and null (pronominal) arguments syntactically and discourse-semantically. The GN operator is located in SAPP, and can license generic inclusive null pronominals semantically:[27]

(47) [CP *Mielőtt kiszáll-unk* proGN *az autó-ból,* [CPellenőriz-zük
    before   get_out-1PL (we.NOM) the car-ELA check_out-IMP.1PL
    proGN, *hogy*   [CP *nem halad-e   gépjármű*
    (we.NOM) whether not   pass.3SG-QCL vehicle
    proGN *mögöttünk*]].
    (we)   behind.POSS1PL
    'Before we get out of the car, we check out whether there is any vehicle behind us.'

---

26 See Ladusaw (1996) for a similar proposal on licensing NPIs.
27 On licensing 3SG generic inclusive lexical and null pronominals see Dalmi (2013, 2014, 2017, 2018).

(48) [_CP_ *Amit   ma   megte-het-sz* pro_GN_], [_CP_ *ne   halasz-d*   pro_GN_
       what   today   do-POT-2SG (you)   not   put_off-IMP.2SG (you)
       holnap-ra].
       tomorrow-ILL
       'Do not put off till tomorrow what you can do today.'

1PL and 2SG null pronominals can be interpreted as generic in (48) and (49) only because they are licensed both syntactically and semantically, in the sense of Ladusaw (1996).

4 Conclusion

This chapter investigated the distribution, licensing conditions and coreference options of lexical and null pronominals in Hungarian matrix and dependent clauses. While individual reference lexical pronominals require merely syntactic licensing via cyclic Agree, null pronominals need to be, in addition, semantically licensed, i.e. they must be in the scope of an operator, located in the Operator Field of the C-domain.

Generic inclusive lexical arguments occur freely in matrix, argument and adjunct clauses, while their null counterparts need to be co-indexed with a generic inclusive antecedent in an adjacent clause and must be in the scope of GN. Generic exclusive lexical and null arguments are free in matrix and dependent clauses alike.

The coreferential interpretation of individual reference lexical pronominals is more restricted than that of null pronominals. Null pronominal arguments can occur in matrix and dependent clauses alike, providing that they have a lexical antecedent in an adjacent clause. Their semantic licensing is ensured via being in the scope of some kind of operator in the C-domain.

References

Abney, Steven. 1986. The English noun phrase and its sentential aspect. Cambridge, MA: MIT PhD dissertation.

Ackema, Peter & Ad Neeleman. 2018. *Features of Person. From the Inventory of Persons to their Morphological Realization*. Cambridge, MA: MIT Press.

Bárány, András. 2015. *Differential object marking in Hungarian and the morpho-syntax of case and agreement*. PhD dissertation. University of Cambridge, Cambridge.

Barbosa, Pilar. 2009. Two kinds of subject *pro*. *Studia Linguistica* 63 (1). 2–58.

Barbosa, Pilar. 2019. *Pro* as a minimal nP: Towards a unified apporach to *pro*-drop. *Linguistic Inquiry* 50 (2). 1–40.

Bartos, Huba. 1997. On objective and subjective agreement in Hungarian. *Acta Linguistica* 44. 262–384.

Bartos, Huba.1999. Morfoszintaxis és interpretáció. Az inflexiós jelenségek szintaktikai háttere [Morpho-syntax and interpretation. The syntactic background of inflectional phenomena.] PhD dissertation. Eötvös Loránd University, Budapest.

Bartos, Huba. 2000. Az inflexiós jelenségek szintaktikai háttere [The syntactic background of inflectional phenomena]. In Kiefer, Ferenc (ed.) *Strukturális magyar nyelvtan 4. Morfológia* [The Syntactic Structure of Hungarian 4. Morphology], 653–761. Budapest: Akadémiai Kiadó.

Bèjar, Susan & Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40. 35–73.

Benedicto, Elena. 1995. *Mne n'e čitaetsja*: relativized modality, datives and reflexive suffixes *UPenn Working Papers* 2. 1–17.

Bianchi, Valentina. 2006. The syntax of personal arguments. *Lingua* 116. 2023–2067.

Bizzarri, Camilla. 2015. Russian as a partial *pro*-drop language. Data and analysis from a new study. *Annali di Ca'Foscari* 49. 335–368.

Bródy, Mihály. 1990. Remarks on the order of elements in the Hungarian focus field. In Kenesei, István (ed.) *Approaches to Hungarian* 3. 95–123.

Bródy, Mihály. 1995. Focus and checking theory. *Approaches to Hungarian* 5. 29–45.

Bródy, Mihály. 2013. Silent *people*. The pseudo-impersonal. In Johan Brandtler, Valéria Molnár & Christer Platzack (eds.), *Approaches to Hungarian 13: Papers from the 2011 Lund Conference* 33–44. Amsterdam: Benjamins.

Cabredo-Hofherr, Patricia 2003. Arbitrary readings of third person plural pronominals. In Matthias Weisgerber (ed.), *Proceedings of the conference "sub7-Sinn und Bedeutung"*, 81–94. University of Konstanz.

Cardinaletti, Anna. 1997. Subjects and clause structure. In Lilianne Haegeman (ed.), *The New Comparative Syntax*, 33–64. London: Longman.

Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In Luigi Rizzi (ed.), *The Structure of CP and IP. The Cartography of Syntactic Structures. Vol 2*. 115–165. Oxford: Oxford University Press.

Cardinaletti, Anna & Ian Roberts. 2002. Clause structure and X-second. In Guglielmo Cinque (ed.), *Functional Structure in DP and IP. The Cartography of Syntactic Structures Vol 1*. 123–166. Oxford: Oxford University Press.

Chierchia, Gennaro. 1995. The variability of impersonal subjects. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Partee (eds.), *Quantification in Natural Languages,* 107–143. Dordrecht: Kluwer.

Cinque, Guglielmo 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19. 521–581.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cognola, Federica & Jan Casalicchio (eds.). 2018. *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*. Oxford: Oxford University Press.

Coppock, Elisabeth. 2013. A semantic solution to the problem of Hungarian object agreement. *Natural Language Semantics.* 21 (4). 345–371.

Coppock, Elisabeth. & Stephen Wechsler. 2012. The objective conjugation in Hungarian: Agreement without *phi* features. *Natural Language and Linguistic Theory* 30. 699–740.

Cysouw, Michael. 2008. Inclusive/exclusive distinction in independent pronouns. In Martin Haspelmath, Matt Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures Online*. Chapter 39. Munich Digital Library.

D'Alessandro, Roberta. 2007. *Impersonal Si Constructions: Agreement and Interpretation*. Berlin: Mouton De Gruyter.

D'Alessandro, Roberta & Artemis Alexiadou. 2003. Inclusive and exclusive impersonal pronouns. A feature-geometrical analysis. *Rivista di Grammatica Generativa* 27. 31–44.

Dalmi, Gréte. 1994. *Hungarian Infinitival Constructions*. Sydney: University of Sydney MPhil thesis.

Dalmi, Gréte. 1998. On object agreement in Hungarian. *Working Papers in the Theory of Grammar*. 5 (2). 91–119.

Dalmi, Gréte. 2000. The structure of *psych*-predicates in Hungarian. *Proceedings of SCL18*. 193–206.

Dalmi, Gréte. 2002. The role of AGRP in non-finite predication. Budapest, Eötvös Loránd University PhD dissertation.

Dalmi, Gréte. 2010. *Copular Sentences, Predication, and Cyclic Agree*. Saarbrücken: Lambert Academic Publishing (VDM Verlag).

Dalmi, Gréte. 2013. ALL-IN-ONE: Generic inclusive subjects in Hungarian. *Proceedings of WCCFL 31*. 115–123. Somerville: Cascadilla Proceedings Project.

Dalmi, Gréte. 2014. The feature geometry of generic inclusive null DPs in Hungarian. In Anna Bondaruk, Gréte Dalmi, & Alexander Grosu (eds.), *Advances in the Syntax of DPs. Structure, Agreement and Case*. Linguistik Aktuell 217, 165–193. Amsterdam: John Benjamins.

Dalmi, Gréte. 2017. Little *pro*'s but how many of them? On 3SG null pronominals in Hungarian. *LingBaW Journal* 3. 61–73.

Dalmi, Gréte. 2018. Az ember és társai, avagy, következetes üres alanyt engedélyező nyelv-e a magyar? [One & Co, or, is Hungarian a consistent null subject language?] *Nyelvtudományi Közlemények* [Linguistic Announcements] 114. 29–43.

Darlymple, Mary & Irina Nikolaeva. 2011. *Objects and Information Structure*. Cambridge: Cambridge University Press.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33 (3). 409–442.

Dvořak, Vera. 2015. Syntax of generic null objects revisited. In *Proceedings of the Annual Meeting of BLS* 41. 71–99.

Egerland, Verner. 2003. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax* 71. 75–102.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.

É. Kiss, Katalin. 2002. *Hungarian Syntax*. Cambridge: Cambridge University Press.

Farkas, Donka. 1987. DO *pro*-drop in Hungarian. In Kenesei, István (ed.), *Approaches to Hungarian* 2. 191–212.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistic Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null-subject languages. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: A synchronic and diachronic perspective*, 211–239. Oxford: Oxford University Press.

Harley, Heidi & Elisabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78 (3). 482–526.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. The null generic subject pronoun in Finnish: A case of incorporation to T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 200–231. Cambridge: Cambridge University Press.

Holmberg, Anders & Michelle Sheehan. 2010. Control into finite clauses in partial null subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–153. Cambridge: Cambridge University Press.

Holmberg, Anders & On-Usa Phimsawat. 2015. Generic pronouns and *phi* features. Evidence from Thai. In Alison Biggs, Man Li, Aiquing Wang & Cong Zhang (eds.), *Proceedings of the Second Asian and European Linguistic Conference*. *Newcastle and Northumbria Working Papers in Linguistics* 21 (1). 55–72. Newcastle: University of Newcastle Press.

Huang, James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15. 531–574.

Jaeggli Osvaldo & Kenneth Safir. 1989. *The Null Subject Parameter*. Dordrecht: Foris.

Jónsson, Jóhannes Gisli. 1997. *Clausal Architecture and Case in Icelandic*. PhD dissertation. University of Amherst, Amherst, MA.

Kenesei, István. 1989. On pronominal binding in Hungarian. In László Marácz & Pieter Muysken (eds.), *Configurationality. The Typology of Asymmetries*. 223–236. Dordrecht: Foris.

Kratzer, Angelika. 2000. German impersonal pronouns and logophoricity. Paper presented at the *Generic Pronouns and Logophoricity Conference*. Sao Paolo.

Krifka, Manfred, Francis J. Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: An introduction. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The Generic Book*, 1–124. Chicago: University of Chicago Press.

Krzek, Małgorzata. 2012. *The Syntax of Impersonal Constructions in Polish*. PhD dissertation. University of Newcastle, Newcastle.

Ladusaw, William. 1996. Negative concord and 'mode of judgment'. In Heinrich Wansing (ed.), *Negation: A Notion in Focus*. 127–143. Berlin: Mouton de Gruyter.

Longobardi, Giuseppe 1994. Reference and proper names. A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25: 609–665.

May, Robert 1977. *The Grammar of Quantification*. PhD dissertation. MIT, Cambridge, MA.

Moltmann, Friderike 2006. Generic one, arbitrary PRO, and the first person. *Natural Language Semantics* 14. 257–281.

Moltmann, Friderike. 2010. Generalizing detached self-reference and the semantics of generic *one*. *Mind and Language* 25 (4). 440–473.

Nikolaeva, Irina. 1999. Object agreement, grammatical realtiosn and information structure. *Studies in Language* 23. 331–376.

Partee, Barbara. 1988. Copular inversion puzzles in English and Russian. In Katarzyna Dziwirek, Herbert S. Coats & Cynthia Vakareliyska (eds.), *Formal Approaches to Slavic Linguistics*. 361–395. Ann Arbor: Michigan Slavic Publications.

Pesetsky, David. 1987. Binding problems with experiencer verbs. *Linguistic Inquiry* 18. 126–140.

Rapoport, Tova. 1995. Specificity, objects and nominal small clauses. In Anna Cardinaletti & Maria-Teresa Guasti (eds.), *Small Clauses*. Syntax and Semantics 28, 153–177. New York. Academic Press.

Ritter, Elisabeth & Martina Wiltshko. 2019. Nominal speech act structure. Evidence from the structural deficiency of impersonal pronouns. *Canadian Journal of Linguistics* 64 (4). 702–729.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Cambridge, MA: MIT Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan. *Parametric Variation: Null Subjects in Minimalist Theory*, 1–58. Cambridge: Cambridge University Press.

Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense and speech features. *Italian Journal of Linguistics* 16. 219–251. (Special issue edited by Valentina Bianchi & Kenneth Safir).

Sigurdhson, H. A. & V. Egerland 2009. Impersonal null subjects in Icelandic and elsewhere. Studia Linguistica 63: 158–185.

Simon, Horst 2005. Only *you*? In Elena Filimonova (ed.), *Clusivity. Typology and Case Studies in the Inclusive / Exclusive Distinction*,113–150. Amsterdam: John Benjamins.

Surányi, Balázs. 2002. *Multiple Operator Movements in Hungarian*. PhD dissertation. University of Utrecht, Utrecht.

Surányi, Balázs. 2004. The left periphery and cyclic spell-out: The case of Hungarian. In David Adger, Cecil de Cat & George Tsoulas (eds.)*, Peripheries. Syntactic Edges and Their Effects*. 49–75. Dordrecht: Kluwer.

Szabolcsi, Anna. 1981. The semantics of Topic-Focus articulation. In Jeroen Groenendijk, Martin Stockhof & Theo Janssen (eds.), *Formal Methods in the Study of Language*, 513–541. Mathematic Centrum Amsterdam.

Szabolcsi, Anna. 1994. The noun phrase. In Ference Kiefer & Katalin É. Kiss (eds.), *The Syntactic Structure of Hungarian*. Syntax and semantics 27, 179–274. New York: Academic Press.

Szécsényi, Krisztina. 2009. *An LF-driven Theory of Scrambling in Hungarian Infinitival Constructions*. PhD dissertation. University of Szeged, Szeged.

Tóth, Ildikó. 2011. Non-referential readings of null subjects in Hungarian. In Tibor Laczkó, György Rákosi & Catherine Ringen (eds.), *Approaches to Hungarian* 12, 209–238. Amsterdam: John Benjamins.

Tsedryk, Egor. 2015. Deriving null pronouns: a unified analysis of subject drop in Russian. In Małgorzata Szajbel-Keck, Roslyn Burns & Darya Kavitskaya (eds.), *Formal Approaches to Slavic Linguistics 23: The First Berkeley Meeting*, 342–361. Ann Arbor: Michigan Slavica Publication.

Jeremy Bradley and Johannes Hirvonen

10 Null subjects in Mari

1 Introductory remarks

Mari[1] is generally described as a *pro*-drop language in reference materials (e.g. Riese et al. 2019: 137), though there is currently little reliable data on the parameters governing the inclusion or omission of subject pronouns in the literature. In the chapter at hand we examine and eventually corroborate the status of Mari as a consistent Null Subject Language (NSL) following established literature (Holmberg 2010; Roberts & Holmberg 2010; D'Alessandro 2015; Frascarelli 2018), through empirical research, i.e. a corpus-based survey using the demo version of the annotated corpus of literary Mari[2] published at corpus.mari-language.com (henceforth *Onchyko* Corpus), an online survey filled out by 39 native speakers of Mari, and elicitations with and grammaticality judgements by native speakers of Mari.

When examining Mari's status as an NSL and its idiosyncratic facets in comparison with other NSLs, the following factors require special attention. (1) How does syncretism influence the verbalization or omission of an overt subject? The Mari verbal paradigm has little syncretism in general, but isolated syncretic forms do exist (e.g. *tol'ǝč* come.PST1.2SG 'you came' ~ *tol'ǝč* come.PST1.3PL 'they came') and the influence of this on pronoun expression calls for investigation. (2) Both 2SG and 3PL are used generically; factors governing referential and generic

2 At the time of our research, the demo version of the corpus contained 3,010,327 tokens of literary Mari taken from the 2012–2014 editions of the literary journal Onchyko which were used in this paper as a small but representative sample of contemporary literary Mari. In December 2020, a corpus containing 57.38 million tokens was published.

readings in these persons are thus examined. (3) The conditions on antecedence for null 3SG pronouns are analysed, as are Topic Chains and null subjects of non-finite/nominalized forms.

2 Basic facts about null subjects in Mari

2.1 Simple clauses

In simple clauses, referential pronominal subjects of the first and second person singular and plural can be dropped without restriction. This is not limited by clause type, tense, or mood. The pronouns can be dropped in affirmative clauses (1)/(5), *wh*-questions (2), negated clauses (3), or exclamative clauses (4).[3]

(1) *məj* / pro[4]   *čaj-əm   šolt-em*[5]
    1SG / (1SG)   tea-ACC   boil-1SG
    'I'm making tea.'

(2) *təj* / pro   *kušto   il-et?*
    2SG / (2SG)   where   live-2SG
    'Where do you live?'

(3) *təj* / pro   *čən-žə-m   og-əda   pale*
    2SG / (2SG)   truth-POSS.3SG-ACC   NEG-2PL   know.CNG
    'You don't know the truth.'

(4) *oχ,   kuze   məj* / pro   *noj-en-am,   koloj-em!*
    oh   how   1SG / (1SG)   become.tired-PST2-1SG   little.child-POSS.1SG
    'Oh, how tired I have become, my little child!'

---

3 Mari has two morphologically formed past tenses, glossed here as PST1 and PST2. The choice of past tense is governed by a conspiracy of factors such as evidentiality, time depth, and aspect (cf. Riese et al. 2019: 220–221). As there is some syncretism in PST1 but none in PST2, the distinction between these past tenses will be indicated in this chapter.
4 In this paper pro is used to refer to a null (i.e. unpronounced) pronoun, in compliance with the convention set in this volume.
5 If not indicated otherwise, the language data in this chapter is in literary Meadow Mari and was taken from the corpus or was produced by our native speaker consultants.

(5) *məskəltəš-əm me* / pro   *tuleč*   *vara*   *čəlt*   *mondə-š-na*
    mockery-ACC   1PL / (1PL)   from.this   after   completely   forget-PST1-1PL
    'We completely forgot the mockery after this.'

Pronominal subjects of the third person singular and plural can be freely dropped as well and do not require an immediate antecedent. The conditions on the interpretation of null and overt 3SG pronouns will be discussed in section 4. The generic interpretation of 3PL pronominal subjects (which can in some cases preclude the omission of a pronominal subject) will be dealt with in section 3.

(6) *tudo* / pro   *kol-əm*   *kuč-əš*
    3SG / (3SG)   fish-ACC   catch-PST1.3SG
    'S/he caught a fish.'

(7) *nuno* / pro   *naverne,*   *kočəvüd*   *dene*   *tol-ən-ət*
    3PL / (3PL)   probably   vodka   with   come-PST2-3PL
    'They probably came with vodka.'

2.2 Expletives

The usage of an overt expletive pronominal subject (cf. English **It** *rains*) cannot be observed in Mari. General circumstances are expressed with a null 3SG subject; the verb *lij-* 'to be, to become; to be allowed, to be possible' is especially common in this usage.

(8) *tek*   *təge*   *lij-eš*
    may   like.this   be(come)-3SG
    'May it be like this.'

Verbs denoting weather-phenomena such as *lum-* 'to snow' and *jür-* 'to rain' typically occur in conjunction with a nominal subject, as illustrated in (9).

(9) a. *lum*   *lum-eš*
      snow   snow-3SG
      'It's snowing.'
   b. *jür*   *jür-eš*
      rain   rain-3SG
      'It is raining.'

There are a number of productive patterns with derivational suffixes in which no overt subject is permissible. Thus, (10) demonstrates the use of the productive denominal verbal suffix *-aŋ*, primarily used to indicate that the quality denoted by a noun has been acquired (cf. Alhoniemi 1985: 161; Riese et al. 2019: 388). In turn, (11) presents the valency-reducing suffix *-alt/-əlt* (cf. Alhoniemi 1985: 160; Pengitov, Galkin & Isanbayev 1961: 161; Riese et al. 2019: 391) yielding the impersonal interpretation, potentially in combination with accusative objects. Finally, in (12) the productive causative suffix *-kt* (cf. Alhoniemi 1985: 163; Riese et al. 2019: 392–393) is used. In all the three constructions exemplified, an overt subject is not permissible.

(10) *urem-əšte   rümbək-aŋ-e*
      street-INE   dusk-VB-PST1.3SG
      'It grew dark outside.'

(11) *ala-kunam   ožno   suas   jəlmə-m=at   tunem-alt-ən*
      sometime   earlier Tatar language-ACC=and study-DTR-PST2.3SG
      'Sometime in the past, Tatar was studied as well.'[6]

(12) *ulo   mogər-žə-m   kəlmə-kt-en*
      Whole   body-POSS.3SG-ACC freeze-CAUS-PST2.3SG
      'S/he felt chills all over his/her body' (lit. 'It froze his/her body.')

The latter construction constitutes a so-called *quasi-causative* construction (cf. F. Gulyás 2013: 38) that can be found in different branches of the Uralic language family (e.g. Finnish, Permic), and notably also in Turkic languages of the Volga-Kama region.

2.3 Verbal syncretism and null subjects

Previous research on NSLs has confirmed that rich agreement on the verb is the main predictive factor for licensing null subjects (Taraldsen 1980; Roberts & Holmberg 2010; D'Alessandro 2015). Generally, languages tend to allow referential null subjects if their verbal agreement paradigms are largely non-syncretic and allow

---

6 Notably, this construction is not permissible in Hill Mari. A potential contact linguistic relevance of this dialectal split requires further investigation.

the null pronoun to be identified by verbal agreement morphology alone, while languages with a larger amount of syncretism do not. This generalization makes general typological and diachronic predictions but does not necessarily predict whether a given language ultimately allows null subjects or not (Gilligan 1987). Typologically, it is more common for verbal paradigms not to neatly fit the dichotomy of fully syncretic or non-syncretic, but to exhibit partially syncretic verbal paradigms where agreement is underspecified and thus syncretic only for some grammatical categories like tense, or only single person-number combinations. There are differences as to whether these languages allow null subjects in both syncretic and non-syncretic environments. For example, Rosenkvist (2010: 237) reports that in Övdalian (Sweden), null subjects are only permissible in the first and second person plural, and these indeed are the only forms in the verbal paradigm with clearly distinct forms. Likewise, whenever the subject cannot be unambiguously identified by verbal agreement, an overt pronoun must be used. The other type of language is exemplified by Mari where null subjects are permissible both in syncretic and non-syncretic environments.

As is typical for consistent NSLs, Mari has rich verbal agreement. All Mari verbs fall into one of two conjugation classes with distinct inflectional endings. The agreement paradigm (two conjugations with three synthetically formed tenses and two moods, negation expressed through a negation verb combined with a connegative form) is largely non-syncretic with only two instances of syncretism, both in the affirmative paradigm (there is no syncretism in negation). Thus, in the first conjugation, 2sg and 3pl forms in the Simple Past I are syncretic (Table 1), while in the second conjugation, 3pl forms of Simple Past I and the Imperative are syncretic (Table 2). Only in the first case does syncretism pertain to person marking; syncretism is not constrained to a grammatical category like a certain tense or mood but only to single forms and cannot be viewed as systematic. The present tense forms of both conjugations show that agreement is mostly non-syncretic.[7]

In spite of the syncretism, null subjects are permissible in 2sg and 3pl in the Simple Past I of the first conjugation:

---

7 See Alhoniemi (1985) or Riese et al. (2019) for a comprehensive overview of Mari verbal morphology; verbal paradigms can be generated at paradigm.mari-language.com. The paradigms of the two verbs used for illustrative purposes here can be produced with the input *nalam* 'I take' and *yshtem* 'I do'.

Table 1: Present, Simple Past I, and Imperative forms of *nal-* 'to take' (first conjugation).

	present	*Simple Past I*	*Imperative*
1SG	nalam	nal'əm	–
2SG	nalat	███████████████	nal
3SG	naleš	nal'e	nalže
1PL	naləna	nalna	–
2PL	naləda	nalda	nalza
3PL	nalət	███████████████	naləšt

Table 2: Present, Simple Past I, and Imperative forms of *əšte-* 'to do' (second conjugation).

	Present	*Simple Past I*	*Imperative*
1SG	əštem	əštəšəm	–
2SG	əštet	əštəšəč	əšte
3SG	əšta	əštəš	əštəže
1PL	əštena	əštəšna	–
2PL	əšteda	əštəšta	əštəza
3PL	əštat	███████████████████████	

(13) *təj* / pro   ~ *nuno* / pro   *serəš-əm   nal'əč?*
      2SG / (2SG)   ~ 3PL / (3PL)   letter-ACC   take.PST1.2SG~3PL
      'Did you/they get the letter?'

(14) *təj* / pro   ~ *nuno* / pro   *memnan   jal-əške   il-aš   tol'əč.*
      2SG / (2SG)   ~ 3PL / (3PL)   2PL.GEN   village-ILL   live-INF   come.PST1.2SG~3PL
      'You/they came to live in our village.'

Thus, Mari allows null subjects in both syncretic and non-syncretic contexts. It nevertheless deserves investigation if any tendencies with respect to the frequency of null subjects in dependence of the verbal syncretism can be observed. The *Onchyko* Corpus can provide preliminary insights. As disambiguation of the corpus is currently underdeveloped, homonymous forms add uncertainty to the search results and can render them meaningless if one does not correct for them. This pertains to the forms under consideration: in the first conjugation, one cannot meaningfully distinguish between the syncretic PST1.2SG and PST1.3PL

forms, while in the second conjugation, one cannot meaningfully distinguish between PST1.3PL and IMP.3PL forms (see Table 3).

Table 3: Syncretic endings, both conjugations.

	Conj. I	Conj. II
PST1.2SG		-šəč
PST1.3PL		
IMP.3PL	-št	

Furthermore, homonymy of forms derived from different stems can complicate the picture, e.g. first conjugation *tol-* 'to come' > *toləšt* 'they should come' ~ second conjugation *tolo-* 'to steal' > *toləšt* 'they stole', *toləšt* 'they should steal'. Reliable data can only be produced if we a) restrict ourselves to looking at verbs not subject to homonymy beyond the syncretism under investigation, b) create aggregate totals of PST1.2SG, PST1.3PL, and IMP.3PL, and compare the frequency of overt subjects.

Highly frequent Mari verbs, not subject to any additional relevant homonymy in the forms under consideration, include the following:

First conjugation: *lij-* 'to be(come)', *nal-* 'to take; to get; to buy', *už-* 'to see', *mu-* 'to find', *tüŋal-* 'to begin', *lekt-* 'to leave', *man-* 'to say', *šogal-* 'to stand up', *piž-* 'to stick', *lukt-* 'to extract', *vašlij-* 'to meet', *oškəl-* 'to step', *kudal-* 'to ride'

Second conjugation: *əšte-* 'to do', *šono-* 'to think', *šənde-* 'to place', *kaje-* 'to go', *mondo-* 'to forget', *pəšte-* 'to put', *šogo-* 'to stand', *uməlo-* 'to understand', *šukto-* 'to carry out', *kalase-* 'to say', *kolto-* 'to send', *pətare-* 'to end'

Table 4 shows how often these forms can be found in total, how often they are preceded by a 2SG or 3PL pronominal subject in the sentence (pragmatically neutral word order), and how often they are immediately followed by a pronominal subject (pragmatically marked word order).

Table 4: Frequency of pronominal subject in PST1.2SG, PST1.3PL, and IMP.3PL.

	Conj. I test set	Conj. II test set
Total	2230	1443
Preceded by pron. subj.	199 (8.92 %)	84 (5.82 %)
Immediately followed by pron. subj.	2 (0.09 %)	1 (0.07 %)
Overt pronominal subject (total)	201 (9.01 %)	85 (5.89 %)

The larger share of overt pronouns in the first conjugation, where person-based syncretism occurs, is statistically highly significant.[8] These preliminary results suggest that indeed, syncretism of single paradigm cells (and not a regular, system-wide syncretism) has an influence on the usage frequency of null pronouns, and that a null pronoun is more readily available in non-syncretic environments.[9] Nevertheless, even in syncretic environments, null subjects are the norm rather than the exception.

3 Generic readings

One diagnostic distinguishing consistent NSL from partial NSL is the interpretation of the null 3SG pronoun. In partial NSL it can receive a generic interpretation while in consistent NSL only a referential interpretation is possible (Holmberg 2010). Mari mostly corresponds to the consistent NSL type; an overt 3SG pronoun cannot receive a generic reading, with 100% of the informants agreeing on this matter (15), while only 69.2% of the informants preferred a referential interpretation of null 3SG pronouns (16):

(15) *tudo   təšte   sajən   mal-a*
      3SG   here   well   sleep-3SG
      'S/he sleeps well here.' → 100% referential interpretation

(16) pro   *təšte   sajən   mal-a*
      (3SG)   here   well   sleep-3SG
      'S/he sleeps well here.' → 69.2% referential interpretation

Our native speaker consultant categorically rejected the generic reading of (16) and suggested obtaining the generic reading[10] by using the verb *lij-* 'to be, to become; to be allowed, to be possible' in 3SG. Consider (17):

---

8 p = 0.0006 in a N-1 Chi-squared test, cf. www.medcalc.org/calc/comparison_of_proportions.php.
9 A reviewer points out that other factors might explain the difference in overt pronoun usage, such as semantic differences between the verbs of the two conjugation classes, or the frequent use of a certain construction. This concern is valid, as strong tendencies do exist in this respect, with first-conjugation verbs tendentially associated with intransitivity and second-conjugation verbs tendentially associated with transitivity, esp. in derivation (Riese et al. 2019: 376, 387–396). These tendencies are, however, by no means deterministic and we believe that the test sets we chose are comparable in this regard.
10 Prompted by Russian *Zdeś χorošo spitśa* 'It is good to sleep here.'

(17) *təšte sajən mal-aš   lij-eš*
      here   well   sleep-INF be-3SG
      'One can sleep well here.'

A similar question showed similar preferences (66.7% preferred a referential interpretation in a sentence with a null subject). This disagreement between the respondents on a referential reading already implies a deviation from the norm for a consistent NSL as these generally do not allow generic readings with null 3SG pronouns. We assume this to be a Russian influence on Mari among younger speakers.[11]

For the expression of genericity, Mari resorts to null 3PL and 2SG pronouns. The general distinction is between the inclusive generic reading and the exclusive generic reading. According to Moltmann (2006, 2010), the former is characterized by a reading in which the speaker expresses both a generic statement and at the same time refers to other people by ascribing them properties of the self. Thus, it refers to people in general while including the speaker. The exclusive generic reading refers to people in a domain that exclude the speaker. In Mari, there are tendencies to use 2SG for inclusive generic readings, and 3PL for exclusive generic readings. The following example shows a generic sentence using a 2SG null pronoun in the context of inclusive genericity:

(18) pro   *mo-m   əšt-et, ajdemə-n   natur-žo   tugaje*
      (2SG) what-ACC do-2SG human-GEN nature-POSS.3SG such
      'What're you going to do, that's human nature.'

As is the case for 2SG null pronouns, in the generic usage of 3PL for exclusive genericity a pronominal subject is typically absent, but this is by no means absolute:

(19) Context: You have read that in the near future elections will be held, and you tell this to your friend. What do you say?
      nuno / pro *sajləmaš-əm šukt-aš   šon-at.*
      3PL / (3PL) election-ACC conduct-INF think-3PL
      'They are planning to conduct elections.' → 79.5% with no overt subject

---

11 Note that it made no difference whether Russian translations with the reflexive *-śa* (see above) or with a null pronoun and *možno* 'it is possible' were provided.

(20)   Context: You have read that people in India speak many languages, and you
       tell this to your friend. What do you say?
       Indij-əšte   nuno / pro *šuko   jəlme   dene kutər-at.*
       India-INE   3PL / (3PL) many   language   with   speak-3PL
       'In India they speak many languages.' → 74.4% with no overt subject

Turning to the difference between 2SG and 3PL, we note the tendencies men-
tioned above if the same sentence is presented in different contexts, with the
context suggesting an inclusive or exclusive generic reading: In contexts of inclu-
sive genericity that include the speaker in the speech context, our respondents
showed a preference for 2SG forms, in contrast to contexts of exclusive genericity
that did not include the speaker. Consider the following pair of examples with
different contexts:

(21)   Context: You want to go to the theatre and dress well. Your daughter asks
       you why you dress that way. What do you say?
       pro   *teatr-əš   kaj-et   / kaj-at gən,* pro   *saj   vurgem-əm*
       (2SG/3PL) theatre-ILL go-2SG / go-3PL if   (2SG/3PL) good clothing-ACC
       čij-et   /   čij-at
       put.on-2SG   /   put.on-3PL
       'If you/they go to the theatre, you/they put on good clothes.' → 43.6% 2SG

(22)   Context: You see on television that people are going to the theatre. Your
       daughter asks you why they are dressed well. What do you say?
       pro   *teatr-əš   kaj-et   / kaj-at gən,* pro   *saj   vurgem-əm*
       (2SG/3PL) theatre-ILL go-2SG / go-3PL If   (2SG/3PL) good clothing-ACC
       čij-et   /   čij-at
       put.on-2SG   /   put.on-3PL
       'If you/they go to the theatre, you/they put on good clothes.' → 12.8% 2SG

The context for the second item in (22) does not involve the speaker's putting
on good clothes and thus there is a decrease in 2SG usage. The usage of 3PL is
generally the preferred strategy for a generic statement, and it is preferred even
more strongly in the situation which does not include the speaker. According to
our informant, however, usage of the 2SG null pronoun in the sentences above
allows both a referential reading and, with marked intonation, a reading as a pre-
scriptive rule of conduct, i.e. the exclusive generic interpretation; however, 3PL
is preferred in such contexts. In a different example with clear inclusive generic
interpretation, however, which excludes the prescriptive rule of conduct reading,
usage of 2SG was overall preferred:

(23) Context: In the last weeks, you have often received advertisements without address, and you tell this your sister. What do you say?

južgunam pro *örmašan objavlenij-əm nal-at* / *nal-ət*.
sometimes (2SG/3PL) odd ad-ACC get-2SG / get-3PL
'Sometimes you/they get odd advertisements.' → 74.4% 2SG

Concluding this section, we note that Mari for the most part fits the generalizations about the expression of genericity in consistent NSL, i.e. a null 3SG pronoun is preferably interpreted with a referential reading, null 2SG pronouns are preferred in contexts of inclusive generic readings, and null 3PL pronouns are preferred in contexts of exclusive generic readings, emerging in utterances of prescriptive nature. However, we also note that this picture is not absolute.

4 Finite subordination and the interpretation of null 3SG pronouns

This section deals with the admissibility of null subjects in finite subordinate clauses in general and introduces the conditions for antecedence of null 3SG pronouns in an embedded clause. One diagnostic to differentiate consistent NSLs from partial NSLs is the interpretation for antecedence of 3SG pronouns. In a consistent NSL like Italian, assuming a pragmatically neutral context, a null 3SG pronoun (24a) in an embedded clause indicates coreference, while the overt pronoun in (24b) indicates disjoint reference; in a partial NSL like Finnish, (25), this is optional (cf. Holmberg 2010: 91):

(24) a. *Gianni₁ dice che* pro₁/*₂ *vuole comprare una macchina.*
Gianni says that he wants buy a car
'Gianni₁ says that he₁/*₂ wants to buy a car.'

b. *Gianni₁ dice che lui*₁/₂ *vuole comprare una macchina.*
Gianni says that he wants buy a car
'Gianni₁ says that he*₁/₂ wants to buy a car.'
(Italian, Holmberg 2010: 91)

(25) *Gianni₁ sano-o että* pro₁/₂ *halua-a osta-a auto-n.*
Gianni say-3SG that (3SG) want-3SG buy-INF car-GEN
'Gianni₁ says that he₁/₂ wants to buy a car.'
(Finnish, Holmberg 2010: 91)

This contrast has been attributed to the functions of overt pronouns in consistent NSLs, namely expressing emphasis, focus or topic shift (Cole 2009; Roberts & Holmberg 2010: Section 12).

In Mari, finite subordination with an overt complementizer is rare and for the most part, clausal nominalization is used instead (see Section 6). Especially in the folklore texts of the late 19th and early 20th century that predate strong Russian influence on Mari, finite clauses are often simply juxtaposed without a conjunction, with their semantic relation being communicated from context (Bereczki 1990: 76). Here too, null subjects are admissible, also when the subjects of the two clauses differ.

(26) *məj* / pro   *təlat*   *teŋeče*   *ušeštar-en-am,*   *təj* / pro   *möŋə-štö*
      1SG / (1SG) 2SG.DAT   yesterday   remind-PST2-1SG   2SG / (2SG) home-INE
      oksagalta-t-əm   *mond-en-at.*
      wallet-POSS.2SG-ACC   forget-PST2-2SG
      'I reminded you yesterday that you forgot your wallet at home'

A small set of overt complementizers that couple with finite verb forms and follow the phrases they head are used; most notably *manən* 'that' (lit. 'saying', 27, see below), conditional *gən* 'if' (28), concessive *gənat* 'even if' (< *gən=at* if=and). Much like in simple clauses, null pronouns are permissible here in Mari.

(27) *məj* / pro   *kol-en-am*   *manən*   *te* / pro   *šonə-š-ta*   *čaj*
      1SG / (1SG) die-PST2-1SG   that   2PL / (2PL) think-PST1-2PL   so
      'So did you think I had died?'

(28) *təj* / pro   *ot*   *jü*   *gən,* *məj* / pro   *ške jü-ən*   *kert-am*
      2SG / (2SG) NEG.2SG drink.CNG if   1SG / (1SG) self drink-CVB be.able-1SG
      'If you don't drink, I can drink it myself.'

The complementizer *manən* deserves special attention. It is a grammaticalized form of the converb of *man-* 'to say'. The form *manən* is synchronically used a) adverbially ('saying'), b) as a quotative marker, introducing reported speech and discourse, c) as a complementizer, heading subordinate clauses in general (Toldova & Serdobol'skaya 2014; Riese et al. 2017: 302–303, 354). Clauses using *manən* as a complementizer can modify verbs of diction as well as cognitive and emotive predicates, indicating its grammaticalization as a general complementizer (Toldova & Serdobol'skaya 2014). The strategy of using a converb form of a verb of diction as a complementizer is common in the Volga-Kama Region and throughout Northern Eurasia; cf. Udmurt *šuįsa*, Tatar *dip*, Chuvash *tese*, and equivalent forms in Mongolic and Tungusic languages (Klumpp 2016: 539; Matić & Pakendorf

2013). It is not however found in Komi and the Mordvinic languages, related Uralic languages that were not subject to the same Turkic influence as Mari and Udmurt; *manən* can thus be classified as a Turkic calque (Isanbayev 1961). Turkic languages of the Volga-Kama Region, for example Mishar Tatar, also exhibit a phenomenon known as Indexical Shift (Podobryaev 2014). Indexical Shift describes a phenomenon where the context for the interpretation of an indexical (e.g. *I*, *you*, *here*, *now*) shifts, so that the indexical is no longer interpreted in the utterance context of the matrix clause but in the context of the embedded attitude predicate without syntactically being a quotation (see Deal 2020 for a recent overview). Prima facie indexical shift may be encountered in Mari as well:

(29)   *Anna₁*   pro₁/₂   *mo-m*   *nal-am*   *manən*   *ojl-en?*
        Anna   (1SG)   what-ACC   buy-1SG   that   say-PST2.3SG
        'What did Anna say that I / she was buying?'

The embedded verb *nal-* 'to buy' is in the first person singular, and therefore *pro* must be read as such as well. However, this first person expression can be coreferential either with *Anna* or with the speaker. Syntactically, this sentence is ambiguous between an embedding structure in which the embedded first person expression has undergone indexical shift and so refers to *Anna*, an embedding without indexical shift and coreference of *pro* with the speaker, and a quotation. Our informant confirmed that the embedded parse is available because the *wh*-expression *mom* can take scope over the matrix clause, which would be excluded in a quotation structure (cf. Shklovsky & Sudo 2014). The interpretations of this sentence were indeed mixed. Our informants preferred the reading in which the null pronoun referred to *Anna* (64.1%), i.e. either a quotation reading or a shifted reading. When an overt 1SG pronoun was used, however, the preferred reading changed:

(30)   *Anna₁*   *məj₁/₂*   *mo-m*   *nal-am*   *manən*   *ojl-en?*
        Anna   1SG   what-ACC   buy-1SG   that   say-PST2.3SG
        'What did Anna say that I/she was buying? / Anna said: "What am I buying?"'

Here, 76.9% of the informants preferred the reading in which *məj* refers not to *Anna* but to the actual speaker.[12] The difference is striking when compared to null pronoun usage. Since the Turkic languages in close vicinity like Mishar

---

12 Note that this sentence is still ambiguous between a quotation and an embedding. In spoken language the disambiguation can be assumed to happen through prosodic means.

Tatar (Podobryaev 2014) or Poshkart Chuvash (Knyazev 2019) do not allow indexical shift with overt pronouns, we tentatively assume that overt pronouns prohibit indexical shift in Mari and that the difference between (29) and (30) can be explained by the unavailability of embedding with indexical shift in the latter case. This must be confirmed by further research, but we can already note a striking interpretative difference between null pronoun and overt pronoun usage here.

This circumstance is relevant for the discussion pertaining to the interpretation of 3SG forms. Indexical shift and the general preference for quotative strategies in *manən*-clauses (Toldova & Serdobol'skaya 2014) may influence the preferred interpretation of null 3SG pronouns within them. Specifically, if a sentence like (29) – in which the matrix subject and embedded subject are coreferential – uses a 1SG null subject under indexical shift or quotation in the *manən*-clause, this may have an effect on the interpretation of embedded null and overt 3SG pronouns in a similar position. Indeed, our informant showed a strong preference for a non-coreferential reading of 3SG and saw a 1SG null subject necessitated by coreferentiality, see the elicited example in (31).

As in (30), an overt subject expression forced the reading of 1SG in reference to the actual speaker, with no indexical shift. Combining these considerations, we tested the interpretation of similar sentences with two possible overt antecedents for the null 1SG pronoun, and for an overt or null 3SG pronoun, in our online survey, see (32).

(31)　a.　pro₁　　pro₂　　*čəla*　　　*əšt-en*　　*kert-am*　　　　*manən*　*šon-a*
　　　　　　(3SG)　(3SG/1SG) everything do-CVB　be.able-1SG　that　　　think-3SG
　　　　　　'S/he thinks that s/he / I (speaker) can do everything.'

　　　b.　pro₁　　pro₂　　*čəla*　　　*əšt-en*　　*kert-eš*　　　*manən*　*šon-a*
　　　　　　(3SG)　(3SG)　everything do-CVB　be.able-3SG　that　　　think-3SG
　　　　　　'S/he₁ thinks that s/he₂ (somebody else) can do everything.'

　　　c.　*tudo*　*məj*　*čəla*　　　*əšt-en*　　*kert-am*　　　*manən*　*šon-a*
　　　　　　3SG　　1SG　everything do-CVB　be.able-1SG　that　　　think-3SG
　　　　　　'S/he thinks that I (speaker) can do everything.'

(32)　*Anna*　*ńi-mo-m*　　　*ojl-en*　　*o-g-əl,*　　　. . .
　　　　Anna　NEG-what-ACC say-CVB　NEG -3SG-be
　　　a.　. . .　*a*　*Julija*　pro　　*kniga-m*　*nal-am*　　*manən*　*ojl-en*
　　　　　　. . .　but　Julija　(1SG)　book-ACC buy-1SG　　that　　say-PST2.3SG
　　　b.　. . .　*a*　*Julija*　pro　　*kniga-m*　*nal-eš*　　*manən*　*ojl-en*
　　　　　　　　　but　Julija　(3SG)　book-ACC buy-3SG　　that　　say-PST2.3SG

c. ... *a   Julija tudo   kniga-m   nal-eš   manən   ojl-en*
      but Julija 3SG   book-ACC buy-3SG that   say-PST2.3SG
'Anna hasn't said anything, but Julija said that (s)he would buy the book.'

When 1SG forms are used without an overt subject (32a), our respondents strongly preferred (85%) the reading in which *pro* refers to *Julija* rather than *Anna* or other people, which is an expected result given the preference for this strategy after speech verbs (Toldova & Serdobol'skaya 2014). If, however, a null pronoun was used with 3SG morphology (32b), only 41.9% preferred a reading in which *Julija* buys the book, while 34.9% preferred *Anna*, and 16.3% another person altogether. Lastly, when an overt 3SG pronoun is used (32c), coreference with *Anna* or another person amounted to 54.8% and 21.4% respectively, while only 16.7% assumed a coreference with *Julija*. These results are mostly in line with observations about the antecedence of 3SG pronouns in consistent NSLs. In this group, an embedded overt 3SG pronoun usually does not allow an interpretation in which it is coreferential with the subject of the main clause (cf. Holmberg 2010), and coreference of *tudo* with *Julija* in example (32c) was indeed ruled out by most participants in our survey. However, there is variation in regard to the reading of null 3SG pronouns since their interpretation depends on the topic status of an antecedent. The results from (32a–c) can tentatively be compared to the results in a survey by Frascarelli (2018: 224). As noted there, the closest antecedent for *pro* in the matrix clause is not the preferred candidate, and a sentence with two antecedents is mostly judged to be ambiguous.

What we do not find in our examples, however, is a strong rejection of coreference of *pro* with *Julija* in (32b), which would be expected if such an interpretation would be ruled out simply by the possibility of expressing the same proposition with a quotation or indexical shift and thus with 1SG morphology as in (32a). We therefore conclude that the preference for the usage of constructions with indexical shift or quotations does not have an effect on the interpretation of 3SG pronouns. The difference between quotations and indexical shift in Mari demands further empirical research (but see Serdobol'skaya & Toldova 2011 and Toldova & Serdobol'skaya 2014).

In summation, we find that speakers prefer to interpret overt 3SG subjects in embedded *manən*-clauses as someone other than the closest antecedent, while null 3SG subjects yield ambiguous readings. 1SG forms with null pronouns in *manən*-clauses express coreference with 3SG matrix subjects, but overt 1SG pronouns refer to the speaker of the utterance. Apart from these facts, the usage of null subjects in embedded clauses can be compared to that in simple clauses.

5 Topic Chains

Next, we turn to the interpretation of null 3SG pronouns in contexts without immediate antecedents. Another difference between partial NSL and consistent NSL is that the former require an antecedent in a higher clause that directly controls 3SG pronouns, while the latter allow a control relation across an intervening subject (Frascarelli 2018: 222–223). For Italian, exemplifying the latter type, it has been shown that the antecedent of a null subject does not necessarily have to be the closest c-commanding antecedent but rather the topic of the preceding clause or the discourse in general (Frascarelli 2007, 2018). (32a–c) above showed a similar tendency for Mari: the closest c-commanding NP is not necessarily the preferred antecedent for a null pronoun.

In this section we first discuss larger scale discourse examples from corpora and then turn to antecedence in a preceding finite clause and the question whether the grammatical role of the antecedent has an effect on the interpretation of null 3SG pronouns.

In contrast to most theories which attribute the licensing of *pro* mostly to morphological factors such as rich agreement, Frascarelli (2007, 2018) links licensing to information structural conditions. She argues that in a consistent NSL null subjects are located in an A'-position in the high Topic field of the C-domain of a clause and that every predicational clause contains such a position, whether overtly filled or not. The basic idea of her proposal is that multiple clauses can refer to the same Aboutness-Shift-Topic (A-Topic) in a Topic Chain, and the continued topic can remain null after it has been introduced. The main distinction to be drawn is the one between A-Topic and Familiar/Given Topic (G-Topic). A-Topics are characterized by denoting what the sentence is 'about' and at the same time being newly introduced to shift the topic of the discourse. G-Topics on the other hand refer to given information in the discourse and can either continue an A-Topic or refer to a constituent as part of the background which is not the A-Topic. Typically, once an A-Topic is established, it remains null in the following C-domains, forming a Topic Chain, until a new A-Topic is introduced. Crucially, only A-Topics break this Topic Chain and require an overt pronoun to continue the previous A-Topic, but G-Topics do not break the chain. We tentatively assume that these holds for Mari, as can be observed in the following written example:[13]

---

[13] Example retrieved from газетамарийэл.рф/?p=11962 (accessed 10 February 2021).

(33)   ***Avam*** *pašam motkočak jöraten.* **pro** *tugaj poro-poro lijən.* **pro** *ikte denat vursedələn, məskələn, iktən nergenat osaləm ojlen ogəl. A eše, paleda, məjəm* **mo** *en čot örəktaren? Iktaž-mogaj paša əštəšaš lijən gən,* ***tudo*** *verəštəže šinčen-šogen kertən ogəl. Mutlan, pareŋgəm luktaš al'e pum kondaš al'e šudəm əštaš külən gən, avamən čonžo vurgažən. Čerle ulmə**ž** godəmat* **pro** *čəten kertən ogəl.* '**My mother** really loved work. **She** was so very good. **She** would not argue with, mock, or say bad things about anyone. But still, you know **what** surprised me the most? If there was any work to do, **she** couldn't sit or stand around. For example, if potatoes had to be harvested, wood had to be brought in, or hay had to be mowed, my mother (lit. my mother's soul) was worried. Even while **she** was sick, **she** couldn't put up with it.'

In the first sentence, *avam* 'my mother' is the subject in an unmarked position and is introduced as A-Topic. In the following two sentences, a null subject is used with reference to this same continued Topic. Next, however, a new A-Topic, *mo* 'what' is introduced in the position after the object *məjəm* 'me.ACC'. This deviation from the pragmatically neutral word order indicates a contrastive topic, which breaks the Topic Chain, and therefore reference to the mother is taken up by the overt pronoun *tudo* '(s)he' in the next sentence. The next sentence, in which the writer provides an example, does not introduce a new A-Topic into the discourse, since the discourse is still about the mother's behaviour. This sentence can be understood as referring to the background part of the previous sentence. In the last sentence, the mother can again be referred to by a null pronoun, and the 3SG possessive suffix -*ž* in the nominalization *čerle ulməž* 'her sick-being' also refers to her as the subject of the nominalized clause (see Section 6). The null pronoun does not pick up the immediate antecedent *avamən čonžo* 'my mother's soul'. This is different to the overt usage of *tudo* after the contrastive topic; in contrast to that case, the Topic Chain has not been broken since there is no A-Topic in the sentence where the example is provided and therefore no overt pronoun has to be used. It seems that this is rather a tendency than a rule in Mari, however. When asked about whether *tudo* could be left out, our informant regarded its usage mostly as a stylistic choice and thought it typical of literary language, but did not oppose omission of the pronoun. In any case, it is notable that an overt pronoun was used here, but not in the second and third sentence which immediately follow the sentence in which the A-Topic is established. For another similar, more transparent example, consider the following part of the report:[14]

---

[14] Example retrieved from cholman.rbsmi.ru/articles/istoriy-gych-/ilish-mariy/ (accessed 10 February 2021).

(34) *Tidən šotəšto* **kok primerəm** *kondənem.* **Iktəže** *S.S. Novikovən Kirov oblast'*
Malməž kundeməšte tatarla mutlanəše marij-vlakəm **vašlijməže**. **Nuno**
tatarla mutlanat, no škenəštəm marijlan šotlat.
'In this respect I want to present **two examples**. **The first one** is the one
concerning S.S. Novikov's **meeting** with Tatar-speaking Maris in the Malməž
region of the Kirov Oblast. **They** speak Tatar but consider themselves Maris.'

The first sentence introduces a new A-Topic, namely the two examples, of which
the first one is taken up in the following sentence and made the new A-Topic.
Also note the use of the 3SG possessive suffix -*že* in *iktəže* 'the first one (of them)',
which is used in the function of marking the membership of its host in a previously
introduced set, namely the two examples (cf. Simonenko 2014). The Mari (*marij-*
vlak) are introduced in the same sentence as *iktəže* as the accusative object of
a participle being used as a verbal noun (*vašlij-mə-že* meet-PTCP.PASS-POSS.3SG
'his/her meeting'). The overt 3PL pronoun *nuno* is used in the following sentence
to refer to the Mari. In this case our informant claimed that the usage of the overt
pronoun was obligatory and that its omission would render the sentence incom-
prehensible, potentially due to ambiguity with a generic reading.

Frascarelli (2018) claims that Topic Chains can be started by a silent A-Topic.
Concretely, if there is no overt A-Topic in the preceding clause, any constituent
of that clause can function as a silent A-Topic in the next clause. The same can
be observed for Mari. Consider the following question-answer pair, in which the
question introduces a non-subject DP as topic, namely as complement of the
postposition *nergen* 'about':

(35) Q: pro   *Pötər   nergen   iktaž-mo-m   kol-ən-at   mo?*
           (2SG) Pötər   about   INDF-what-ACC   hear-PST2-2SG   Q
            'Have you heard something about Pötər?'
      A: pro / *tudo   paša-m   mu-ən.*
           (3SG) / 3SG   work-ACC   find-PST2.3SG
            'He has found a job.'

In such cases, the use of an overt pronoun in the answer hardly makes any differ-
ence regarding reference: with an overt pronoun 92.3% and with a null pronoun
87.5% of respondents preferred an interpretation in which the second sentence
referred to *Pötər*. We therefore note that both null and overt 3SG pronouns can
refer to a non-subject NP if it is the only overt antecedent. Following the survey
in Frascarelli (2018), we examined whether the grammatical role of an intended
antecedent in a preceding root clause had any effect on the interpretation of the
null pronoun. Consider the following sentence pairs, where there is a nomina-

tive subject, *Jelu* (a feminine name), in the first root clause, and a non-subject NP, *Serge* (a masculine name), in various other grammatical functions: as dative object in (36), as complement of the postposition *dene* 'with' in (37), and again as complement of said postposition, but topicalized in (38).

(36) *Jelu teŋgeče Serge-lan čən-əm ojl-en.   ənde* pro
      Jelu yesterday Serge-DAT truth-ACC tell-PST2.3SG now (3SG)
      saj-ən   mal-en   kert-eš.
      good-ADV   sleep-CVB be.able-3SG
      'Yesterday Jelu told Serge the truth. Now s/he can sleep well.'

(37) *Jelu teŋgeče Serge dene kutər-en.   ənde* pro *saj-ən*
      Jelu yesterday Serge with talk-PST2.3SG now (3SG) good-ADV
      mal-en   kert-eš.
      sleep-CVB be.able-3SG
      'Yesterday Jelu talked with Serge. Now s/he can sleep well.'

(38) *Serge dene Jelu teŋgeče kutər-en.   ənde* pro *saj-ən*
      Serge with Jelu yesterday talk-PST2.3SG now (3SG) good-ADV
      mal-en   kert-eš.
      sleep-CVB be.able-3SG
      'With Serge, Jelu talked yesterday. Now s/he can sleep well.'

In (36) and (37), the preferred reading is coreference of *pro* with *Jelu* (75.6% and 79.1% respectively), and thus the subject and topic in both cases. In (38), where subject and topic do not coincide, however, only 54.3% of our informants chose the reading in which *Jelu* can sleep well, 37% in which *Serge* can sleep well, and 8.7% the one in which someone else can sleep well. This sentence was perceived as ambiguous by 15.2% of participants who chose both *Jelu* and *Serge*. Therefore, the effect of the grammatical role seems to be negligible since there was barely any difference between the preferred interpretation of (36) and (37), but a bigger difference between (37) and (38) which both included the postpositional phrase as a possible antecedent. What seems to play a more important role is the information structural position of the antecedent.

6 Null subjects with non-finite verb forms

In Mari, the finite paradigm is supplemented by a large number of non-finite forms used extensively in clausal subordination. Table 5 illustrates participles

and converbs[15] found in Mari. All participles but the active participle are used as verbal nouns referring to an action as a whole, and not just as verbal adjectives (cf. Riese & al. 2019: 252–253). In this function, they serve as the head of embedded clauses. The various participles and converbs instantiate different syntactic structures, which can be observed by whether a prima facie subject (different from the matrix subject) is possible in the first place, and if it is, whether its *phi* (person and number) features can be cross-referenced by possessive suffixes.

Table 5: Non-finite verbal forms in Mari.[16]

		Suffix	*əšte-* **'to do'**	**Subj-POSS**
Participles	Passive	*-mE*	*əštə-me* 'done'	+
	Fut.-Necessitive	*-šaš*	*əštə-šaš* 'to do/be done'	+
	Negative	*-DəmE*	*əštə-dəme* 'not doing/done'	+
Converbs	Instructive	*-ən/-en*	*əšt-en* 'doing'	−
	Negative	*-De*	*əštə-de* 'not doing'	−
	Anterior	*-mek(e)*	*əštə-mek(e)* 'after doing'	+
	Future	*-meš(ke)*	*əštə-meš(ke)* 'before doing'	+
	Contemporaneous	*-šala*	*əštə-šala* 'while doing'	~

Nominal agreement morphology provides further evidence for the link between agreement and the licensing of null subjects, as postulated by the 'rich agreement hypothesis' (see Section 2.3). The role of agreement for licensing and its connection to the nominal domain was highlighted for example by Gilligan (1987: 165) for Ecuadorian Quechua, which has null subjects in finite matrix clauses where there is verbal agreement (39), but no null subjects in clausal nominalizations which lack agreement (40).

(39) *(ñuka)*   *Marya-ta*   *juya-ni*
      (I.NOM)   Marya-ACC   love-PRS.1SG
      'I love Marya.'
      (Ecuadorian Quechua, Cole 1982: 34)

---

15 An adverbial non-finite form, in other nomenclatures gerunds or adverbial participles.
16 Upper-case *E* indicates vowel-harmonic alternation *e ~ o ~ ö*; uppercase D indicates voice assimilation-based alternation *d ~ t*.

(40) *Juan-ka* [*(*ñuka*) *Marya-ta*   *juya-j-ta*]   *ya-n.*
      Juan-TOP (I.NOM)   Marya-ACC   love-NMLZ-ACC think-PRS.3SG
      'Juan thinks that I love Marya.'
      (Equadorian Quechua, Cole 1982: 34)

According to the hypothesis that null pronouns have to be licensed by sufficiently rich verbal agreement, the pronoun may not be null in combination with this nominalization as there is no agreement morphology that would identify the subject. In the Uralic languages, however, agreement in the nominal domain is widespread, and the subject of many non-finite forms can indeed be marked only by a possessive suffix in Mari. In what follows, we will focus only on the Mari prior and future converbs *-mek(e)* and *-meš(ke)* that can take possessive suffixes of all persons (other non-finite forms are subject to restrictions). For overviews about participles and converbs in Mari, see Pomozi (1997) and Brykina & Aralova (2012).

As is evident from their glossing and name, *-mek(e)* and *-meš(ke)* denote a temporal relation to the matrix clause. The verb retains its argument structure and the converbial clauses correspond to full clauses that can for example have accusative objects and overt subjects. Subjects are tendentially in the nominative when inanimate (*keče lek-meke* sun rise-CVB.ANT 'after the sun came up') and genitive-marked when animate (*Anuš-ən tol-mekə-že* Anuš-GEN come-CVB.ANT-POSS.3SG 'after Anuš came'). Genitive subjects co-occur with a possessive suffix attached to the converb. We classify the subjects of these converbial clauses are genuine subjects, even when genitive marked. The following two examples show converbs in *-mek(e)* with a genitive subject and the argument structure of the verb from which they are derived

(41) *orkestr-ən*   *marij sem-əm*   *šokt-aš*   *tüŋal-mekə-že,*
      orchestra-GEN Mari   melody-ACC play-INF   begin-CVB.ANT-POSS.3SG
      čerkə-se   *ulo*   *kalək*   *memnan dene lekt-ən.*
      church-ADJ whole people 2PL.GEN with leave-PST2.3SG
      'After the orchestra started playing a Mari melody, all the people from the church left with us.'

(42) *Sergej-ən*   *kap* *vij-əm*   *pogə-mekə-že*   *vele, tudo*
      Sergej-GEN   body strength-ACC gather-CVB.ANT-POSS.3SG only 3SG
      laskanrak   *šülalt-en*
      more.easily breathe-PST2.3SG
      'Only after Sergej had regained his bodily strength could he breathe more easily.'

With respect to case and possessive marking, a bidirectional implication can be observed: if and only if the subject is marked with the genitive, it can be cross-referenced by a possessive suffix, and if the subject is marked with the nominative, it cannot be cross-referenced by a possessive suffix. Possessive suffixes indicating the subjects of a converb can also co-occur with pronominal subjects, overt or null:

(43) *[təj-ən   kajə-mek-et],*   pel   ij   gəč   üdər-na
    2SG-GEN go-CVB.ANT-POSS.2SG half year from daughter-POSS.1PL
    šoč-o
    be.born-PST1.3SG
    'After you left, our daughter was born half a year later.'

(44) pro / *məj* [pro *joča-m*   *əštə-mek-em*]   *mo-m-gənat*
    (I) / I   (1SG) child-ACC make-CVB.ANT-POSS.1SG what-ACC-INDF
    šon-en   *mu-am*
    think-CVB   find-1SG
    'After I have the child, I'll think of something.'

The null subject in (44) is clearly comparable to the overt subjects in (41) and (42) above; the differences between the usage of overt and null subjects can be tested in comparison to their usage in finite subordination. When there is neither an overt subject nor possessive marking, the non-finite form lacking subject licensing is generally coreferential with the superordinate clause.

(45) *ik   meŋgə-m oškedə-meke,   məj noj-en-am=at,   šort-aš*
    one km-ACC   walk-CVB.ANT 1SG tire-PST2-1SG=and cry-INF
    tüŋal-ən-am
    begin-PST2-1SG
    'After walking a kilometre, I got tired and started crying.'

For converbs that cannot take possessive suffixes at all (-*n*, -*De*), this is their standard usage.

It should be noted that (at least first and second person) null subjects in a subordinate clause do not necessitate an overt subject in the matrix clause. Complex structures can be found in which no overt subject occurs at all. There seems to be no difference with respect to the interpretation of the subordinate clause, while in the matrix clause, the same factors with respect to overt pronoun usage (emphasis, topic shift) apply.

(46) *bol'nice gəč lek-meke,*   pro / *məj predsedatel' deke kurž-əm*
      hospital from leave-CVB.ANT (1SG) / 1SG chairperson to   run-PST1.1SG
      'After leaving the hospital, I ran to the chairperson'.

For the converbs of anterior and posterior action, the corpus can give insights on the frequency of subject marking strategies (see Table 6). We note that null subjects (rows 3 and 5) are by far the most frequent strategy. We assume that this is a reflection of the usage of null pronouns in unmarked contexts and overt pronouns in marked ones, parallel to what we observed for subjects of finite forms.

Table 6: Subject marking on converbs in Onchyko texts.

	Anterior action *-mek(e)*	Posterior action *-meš(ke)*
1) GEN + POSS marking	1.29%	1.84%
2) GEN only (no POSS)	0.27%	1.13%
3) POSS only (no GEN)	20.02%	32.07%
4) NOM + No POSS	18.25%	20.39%
5) No marking	60.17%	44.57%

To check whether there are differences with respect to overt pronoun usage in non-finite verbal forms, we examined the difference between the usage of null and overt 3SG pronouns in clauses with *-mek(e)* and *-meš(ke)* that precede a matrix clause containing nominal referents. In the sentences with null subjects, converbs without possessive marking were used while in sentences with overt subjects, the converbs were marked with a 3SG possessive suffix. Consider the following sentence pairs:

(47) pro / *tudə-n   pensij-əš   lek-meške   / lek-meškə-že,*
      (3SG) / 3SG-GEN pension-ILL leave-CVB.FUT / leave-CVB.FUT-POSS.3SG
      Anna   Kužener-əšte   il-en
      Anna   Kužener-INE   live-PST2.3SG
      'Before s/he retired, Anna lived in Kužener.'

(48) pro / *tudə-n   Joškar-Ola-ške   tol-meke   /*
      (3SG) / 3SG-GEN Yoshkar-Ola-ILL come-CVB.ANT   /
      tol-mekə-že,   Jelu pialan lij-ən
      come-CVB.ANT-POSS.3SG Jelu happy be(come)-PST2.3SG
      'After s/he came to Yoshkar-Ola, Jelu became happy.'

In both cases, the results were very clear. The majority of informants (92.3% and 82.9% respectively) preferred a reading in which the null subject was coreferential with the matrix subject, with only few respondents preferring a reading with another person. When an overt subject was used, however, the preferred interpretation was coreference of the overt subject with a person differing from the matrix subject (78% and 77.5%, respectively). The results show that, similarly to their usage in finite clauses, overt pronouns are preferably not interpreted with coreference to the closest possible antecedent but to a different referent.

7 Summary

Through the lens of comparative Uralic studies, Mari exhibits no obviously atypical behaviour; it is a consistent NSL in which overt pronominal subjects are admissible but mostly used in special contexts. One such context is syncretism of single paradigm cells where a statistically significant increase in pronoun usage can be observed. Nevertheless, even in syncretic environments, null subjects are the norm rather than the exception. The same tendency can be observed with respect to subjects of non-finite forms; if identifiable by nominal agreement (possessive suffixes), a subject will mostly be null. Overt pronouns are generally used for emphasis or topic shift; in discourse, a null pronoun will typically continue a topic while an overt pronoun will be used for topic shift, and will not refer to its closest antecedent in a matrix clause or preceding clause.

Another difference between overt and null pronoun usage was observed with a special strategy involving subordinate clauses selected by verbs of diction. In particular, for coreference with a 3SG subject, speakers preferably use embedded null subjects, while overt subjects are used to refer to the actual speaker of the utterance. With respect to generic readings, Mari for the most part fits the generalizations for consistent NSL, i.e. both overt and null 3SG pronouns are mostly interpreted with a referential reading, while null 2SG pronouns are preferred in contexts of inclusive genericity, and null 3PL pronouns in contexts of exclusive genericity.

This chapter provided a preliminary outlook on the properties of null subjects in Mari. The patterns found mostly align with established knowledge in respect to other consistent NSLs, but they evade absolute generalizations. Some questions warrant further investigation, including the possibility for generic readings of null 3SG pronouns, the exact semantic division of labour between null 2SG and null 3PL pronouns in generic statements, a precise delineation of indexical shift in Mari, and conditions on the usage of overt pronouns for topic shift.

References

Alhoniemi, Alho. 1985. *Marin kielioppi* [Mari Grammar]. (Apuneuvoja suomalais-ugrilaisten kielten opintoja varten 10). Helsinki: Suomalais-Ugrilainen Seura.

Bereczki, Gábor. 1990. *Chrestomathia Čeremissica*. [A Cheremis reader] Budapest: Tankönyvkiadó.

Bereczki, Gábor. 1992–1994. *Grundzüge der tscheremissischen Sprachgeschichte [Outlines of the history of the Cheremis language (I-II)* (Studia Uralo-Altaica 34, 35). Szeged: University of Szeged.

Bradley, Jeremy. 2016. *Mari Converb Constructions: Productivity and Regional Variance*. Vienna: University of Vienna PhD dissertation.

Brykina, Maria & Natalia Aralova. 2012. Sistemy pričastij v marijskom i permskix jazykax [Systems of participles in Mari and the Permic languages]. In Ariadna Kuznetsova (ed.), Finno-ugorskije jazyki: fragmenty grammatičeskogo opisanija [Finno-Ugric languages: fragments of grammatical description], 476–520. Moscow: Rukopisnyje pamjatniki Drevnej Rusi.

Cole, Melvyn Douglas. 2009. Null subjects: a reanalysis of the data. *Linguistics* 47(3). 559–588.

Cole, Peter. 1982. *Imbabura Quechua* (Lingua Descriptive Studies). Bd. 5. Amsterdam: North-Holland.

D'Alessandro, Roberta. 2015. Null subjects. In Antonio Fábregas, Jaume Mateu & Michael Putnam (eds.), *Contemporary Linguistic Parameters*, 201–226. London: Bloomsbury Academic Publishing.

Deal, Amy Rose. 2020. *A Theory of Indexical Shift: Meaning, Grammar, and Cross-linguistic variation*. Cambridge, MA: MIT Press.

F. Gulyás, Nikolett. 2013. Towards a classification of impersonal constructions in Komi: A functional-typological approach. In Márta Csepregi, Kata Kubínyi & Jari Sivonen (eds.), *Grammatika és kontextus. Új szempontok az uráli nyelvek kutatásában III* [Grammar and context. New aspects in the research of Uralic languages III], 31–49. Budapest: ELTE.

Federal'naja služba gosudarstvennoj statistiki. 2011. Okončatel'nyje itogy vserossiyskoj perepisi naselenija 2010 goda. [Final results of the 2010 All-Russian Population Census].

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistic Theory* 25(4). 691–734.

Frascarelli, Mara. 2018. The interpretation of pro in consistent and partial null-subject languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*, 211–239. Oxford: Oxford University Press.

Gilligan, Gary Martin. 1987. *A Cross-linguistic Approach to the Pro-drop Parameter*. PhD dissertation. University of Southern California, Los Angeles, CA.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Isanbayev, Nikolay 1961. Deepričastija v marijskom jazyke [Converbs in the Mari language]. Yoshkar-Ola: Marijskoje knižnoje izdatel'stvo.

Klumpp, Gerson. 2016. Semantic functions of complementizers in Permic languages. In Kasper Boye & Petar Kehayov (eds.), *Complementizer Semantics in European Languages*, 529–586. Berlin: Mouton de Gruyter.

Knyazev, Mikhail. 2019. Two SAY-complementizers in Poshkart Chuvash: Subject-orientation, logophoricity and indexical shift under verbs of hearing. *Proceedings of WAFL 14*. Cambridge, MA.

Matić, Dejan & Brigitte Pakendorf. 2013. Non-canonical SAY in Siberia: Areal and genealogical patterns. *Studies in Language* 37(2). 356–412.

Moltmann, Friederike. 2006. Generic one, arbitrary PRO, and the first person. *Natural Language Semantics* 14(3). 257–281.

Moltmann, Friederike. 2010. Generalizing detached self-reference and the semantics of generic One. *Mind & Language* 25(4). 440–473.

Pengitov, Nikolay, Ivan Galkin & Nikolay Isanbayev. 1961. *Sovremennyj marijskij jazyk – Morfologija* [Contemporary Mari: Morphology]. Yoshkar-Ola: Marijskoje knižnoje izdatel'stvo.

Podobryaev, Alexander. 2014. *Persons, imposters, and monsters*. PhD dissertation. MIT, Cambridge, MA.

Pomozi, Péter. 1997. *Satzwertige Partizipien auf -mE/-mAš/δâmE im Tscheremissischen: Problematik und Klassifikation.* [Sentence-valued participles in *-mE/-mAš/δâmE in Cheremis: issues and classification*]. (Bibliotheca Cheremissica II). Szombathely: Savariae].(Bibliotheca Ceremissica II). Szombathely: Savariae.

Riese, Timothy, Jeremy Bradley, Monika Schötschel & Tatiana Yefremova. 2019. *Mari (marij jylme): An Essential Grammar for International Learners*. Ms., University of Vienna.

Riese, Timothy, Jeremy Bradley, Emma Yakimova & Galina Krylova. 2017. *Ongay mariy yylme: A Comprehensive Introduction to the Mari Language*. 3.2. Vienna: University of Vienna.

Roberts, Ian & Anders Holmberg. 2010. Introduction: parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Rosenkvist, Henrik. 2010. Null referential subjects in Övdalian. *Nordic Journal of Linguistics* 33 (3). 231–267.

Saarinen, Sirkka. 1997. Borrowed vocabulary in Mari and Udmurt dialects. In Os Hahmo, Sirkka-Liisa, Tette Hofstra, László Honti, Paul van Linde & Osmo Nikkilä (eds), *Finnish-ugrische Sprachen in Kontakt* [Finno-Ugric languages in contact], 191–196. Maastricht: Shaker.

Serdobol'skaya, Natalia & Svetlana Toldova. 2011. Grammaticalization of the verb of speech in Finno-Ugric languages. *Congressus XI. Internationalis Fenno-Ugristarum Piliscsaba, 9–14. VIII. 2010: Pars V. Dissertationes sectionum et symposiorum ad linguisticam*, 285–292. Budapest: Reguly Társaság.

Shklovsky, Kirill & Yasutada Sudo. 2014. The syntax of monsters. *Linguistic Inquiry* 45 (3). 381–402.

Simonenko, Alexandra. 2014. Microvariation in Finno-Ugric possessive markers. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.), *NELS 43: Proceedings of the 43rd Annual Meeting of the North East Linguistic Society, volume 2*, 127–140.

Taraldsen, Knut Tarald. 1980. *On the Nominative Island Condition, Vacuous Application and the That-trace Filter*. Bloomington, IN: Indiana University Linguistics Club.

Toldova, Svetlana & Natalia Serdobol'skaya. 2014. Glagol reči *manaš* v marijskom jayzke: osobennost'i grammatikalizacii [The verb *manaš* in Mari: features of its grammaticalization]. *Voprosi jazykoznanija* 6. 109–134.

Mikael Vinka

11 Two types of null subjects in South Saami

1 Introduction

This chapter presents a study of third person referential null subjects in South Saami, a Finno-Ugric language spoken in central Norway and Sweden by approximately 1000 native speakers (Scheller & Vinka 2016).[1] The main issue in this chapter concerns the contrast between (1a) and (1b), in particular the distinctions in the range of possible antecedents to the null subject in each example:[2]

(1) a. *Læjsa$_1$ vienhti Pöövle$_2$ jeehti* ____$_{1/2/3}$ *bijle-m öösti.*
   Lisa thought Paul said (3SG) car-ACC bought
   'Lisa$_i$ thought that Paul$_2$ said that s/he$_{1/2/3}$ bought a car.'

 b. *Læjsa jeehti Maarja Pöövle-m$_1$ stilli* ____$_1$ *edtji*
   Lisa said Mary Paul-ACC ordered (3SG) WOLL.PST.3SG
   bijle-m åestedh.
   car-ACC buy
   'Lisa$_1$ said that Mary$_2$ ordered Paul$_3$ to buy a car.'

---

2 Throughout the paper I will gloss the auxiliary edtj- as *woll*, following Abusch (1997). This choice is motivated in Section 4.

South Saami is a consistent Null Subject Language (NSL) (see Vinka 2018), and it is well documented that in such languages a null subject is co-indexed with a topic DP (see, for instance, Frascarelli 2007; Grimshaw & Samek-Lodovici 1996; Holmberg 2010b).[3] The multiple embedding in (1a) provides a discourse where there are two potential antecedents to the null subject in the most deeply embedded clause, namely the subject *Pöövle* in the intermediate clause and the subject *Læjsa* in the highest clause. The null subject may also be anteceded by another topic in the immediate linguistic context (see, for instance, Frascarelli 2007; Holmberg 2005, 2010b). (1b) also involves multiple embeddings, and the subject in the most deeply embedded clause is null. However, the referential properties of the null subject in (1b) are sharply distinct from (1a). In (1b), the null subject can refer to the accusative object in the intermediate clause. This friction presents a challenge for any attempt to formulate a unified account of null subjects in South Saami.

Prima facie, there are at least two possible ways to approach the contrast between (1a) and (1b). One could imagine that the null subjects in the two sentences are identical, but other, independent factors influence how co-construal with the antecedent is established. A possible strategy along these lines could be to invoke Landau's (2004) notion of finite control. This is not an unreasonable hypothesis, since the predicate in the intermediate clause in (1b) is the directive control predicate *stilledh* 'order'. However, South Saami differs from, say, Hebrew, since (1b) in fact also permits an overt third person subject pronoun, without affecting the patterns of co-construal. In this regard, the null subject in (1b) resembles what is known from another type of NSL, namely partial NSL (for instance Holmberg 2005, 2010a; Holmberg & Sheehan 2010), which thus represents the other possible approach. Granted that South Saami is genetically related to Finnish, a partial NSL *par excellence*, this is not an unreasonable conjecture. Nevertheless, it brings along the question why the partial NSL pattern does not generalize to (1a), or vice versa, why the consistent NSL pattern of (1a) does not generalize to (1b).

This kind of variation within a specific language was problematic under the classical switchboard conception of parameters that emerged during the early days of the P&P era (Chomsky 1981, 1982). For instance Rizzi's (1982, 1986) Null Subject Parameter provided a catalogue of properties that accompanied the parameter. The parametric cluster associated with null subjects and rich agreement included seemingly unrelated phenomena such as post-verbal subjects and the absence of *that*-trace effects. By parity of reasoning, non-NSLs were expected to differ accordingly. The empirical predictions of the parameter were debated and challenged in a rich body of cross-linguistic research, such as Safir (1985) and

---

3 See also the works by Sigurðsson (2004) and Bianchi (2006).

Huang (1989), to mention a few. The cross-linguistic challenges were accompanied by domain-constrained variations in specific languages, such as fluctuations in binding domains, that could be traced to mood distinctions (Maling 1984). As it became increasingly evident that the wholesale parameter either over- or under-generated, Borer (1984) sowed the seed for a different understanding of variation, which was embraced in Chomsky (1995), namely the idea that "parametric variation is restricted to the lexicon, and insofar as syntactic computation is concerned, to a narrow category of morphological properties, primarily inflectional" (Chomsky 2001:2); the so-called Borer-Chomsky conjecture (Baker 2008).

Against the backdrop of the Borer-Chomsky conjecture, the answer to the question why the null subjects in (1a) and (1b) behave differently, is the following: the heads of the projections that license the null elements are made up of different combinations of features. It is illuminating to consider how the consistent/partial NSL dichotomy is approached. Holmberg (2010b) argues that it rests on two primary ingredients. In consistent NSLs, the third person null pronoun is referentially deficient. This is compensated for by the fact that T hosts an unvalued referential feature. In partial NSLs, T is not equipped with a referential feature, with the consequence that a referential third person must be expressed by an overt pronoun. An interesting characteristic of partial NSLs is the fact that a third person subject pronoun may be null when it occurs in an embedded clause, as pointed out by numerous scholars (Holmberg 2010b; Holmberg & Sheehan 2010; Modesto 2007; Vainikka & Levy 1999), which is illustrated by Brazilian Portuguese in (2).

(2) *O Feco*₁ *disse que a Dani*₂ *acha que* ___*₁/₂ *ganhou na*
  the Feco said that the Dani thinks that (3SG) won in.the
  loto.
  lottery
  (i)   'Feco says that Dani thinks that she (Dani) won the lottery'
  (ii) *'Feco says that Dani thinks that he (Feco) won the lottery'
  (Brazilian Portuguese, Modesto 2006)

In order to capture the contrast between (1a) and (1b), I will argue that a subclass of control verbs, such as *stilledh* 'order' in (1b), has the capacity to select a subjunctive complement. The literature provides ample precedents for this hypothesis, for instance Landau (2004) on Hebrew and Balkan languages, Kempchinsky (2009) on Spanish and Catalan, Spyropoulos (2007) on Greek, Krapova (2001) on Bulgarian, and numerous more. Drawing on Landau (2015) and Shushurin (2017), I will argue that the relevant class of control verbs licenses a participant feature in the left edge domain of the subjunctive complement. More precisely, I claim that this

feature is encoded in the same head as the subjunctive, which in turn is lexicalized by attracting the future auxiliary. The fact that the embedded subject pronoun is obligatorily co-construed with the direct object addressee argument *Pöövle* of the directive verb in (1b), is the result of Agree between the participant feature and the pronoun, edge-linking in the sense of Sigurðsson (2011, 2014a, b). The feature itself is coindexed with a relevant, syntactically local argument in the next clause up. I will furthermore argue that the null subject in (1b) is derived by PF-deletion, in contrast to (1a), which is an incorporated *pro*, in the sense of Holmberg (2010b).

The chapter is organized as follows. Section 2 provides an overview of NSLs, with a special focus on consistent and partial NSLs. Section 3 introduces the essential facts of the distribution of subject pronouns in finite complements of directive and commissive verbs in South Saami. Section 4 addresses the role of the future auxiliary in the complement clauses under consideration, claiming that it has a subjunctive function. The section also shows that subjunctive complements can be selected by directive and commissive control verbs, which, according to Landau (2000, 2004, 2015), form a natural class of control predicates. In section 5, I develop an analysis of the phenomena introduced in the previous sections. While subjunctive complements exhibit similarities with finite control, I dismiss this option and argue that actual control is not involved. Furthermore, I claim that South Saami subjunctive complements are not full CPs, but rather Mood Phrases, along the lines of Giorgi & Pianesi (2004). At the heart of the analysis is the claim that the null subject in a subjunctive complement is a PF-deleted overt pronoun. *De facto* null subjects, I claim, are precluded due to fact that the complement clause is not a CP. Section 6 compares subjunctive complements and unselected subjunctives, adding further support to the proposed analysis. Finally, in section 7, some concluding remarks are given.

2 Null subject languages

The idea that agreement markers on verbs can represent pronominal arguments harks back at least as far as Panini. Panini's system of the mapping of semantic/thematic roles to grammatical functions, kārakas (see Kiparsky & Staal 1969), was well suited to handle certain types of argument drop and it has provided linguistic reasoning with hugely important incitement (Kiparsky 2002: 21). However, Kiparsky notices that Panini did not consider *pro*-drop part of the sentence grammar *per se*, but rather as "a rule-governed discourse process" (Kiparsky 2002: 21), contingent on the presence of agreement morphology. Indeed, Panini's ideas intersect with modern theories in many aspects. Suffice it to say that viewing agreement as a pronominal has a long history in both theoretically and descriptively oriented gram-

matical traditions, as well as in different shapes and eras of modern generative linguistics (Baker 1996; Chomsky 1981; Jelinek 1984, 1998; Perlmutter 1971; Rizzi 1982, 1986; Alexiadou & Anagnostopoulou 1998). The agreement-as-pronoun hypothesis (AAPH) provides an appealing way to explain why some languages do not require an overt subject in finite clauses. However, it is well-known that the simplest conceptions of the AAPH are insufficient. According to Jaeggli & Safir (1989), depending on the nature of agreement in particular languages, the availability of subject drop differs. They suggest a three-way distinction, in which the presence of rich agreement entails null subjects, whereas weak agreement correlates with the absence of *pro*-drop. Furthermore, drawing on Huang (1989), the third variety is represented by Chinese-style languages, which radically lack agreement, but nonetheless allow argument drop. A related issue is found in Germanic-style topic drop phenomena (Neeleman & Szendrői 2007; Sigurðsson 1993; Sigurðsson & Maling 2008).

However, one of the major theoretical problems with Jaeggli & Safir's characterization lies in the notoriously elusive distinction between strong and weak agreement.[4] It is interesting in this context to notice that one of the most notable champions of the so-called "rich agreement hypothesis", namely Rohrbacher (1999), pointed out that morphology alone does not suffice to account for the different *pro*-drop properties in European and Brazilian Portuguese. The former enforces referential third person *pro*-drop in certain environments in contrast to the latter, in spite of the fact that the inflectional paradigms in the two varieties of Portuguese are equally rich in terms of morphological complexity.[5] In the words of Holmberg (2005), European Portuguese is a consistent NSL, whereas Brazilian Portuguese is a partial NSL.

2.1 Consistent and partial NSLs

I now turn my attention to some of the details that distinguish consistent NSLs and partial NSLs, and in the process I will show that South Saami is a consistent NSL.

The distribution of overt and covert third person subject pronouns in consistent NSLs is contingent on whether or not the pronoun refers to a topic (Frascarelli

---

4 While Chomsky's (1995: ch. 2) introduction of strong/weak features into the theory was highly influenced by Pollock's (1989) seminal work on verb movement, the features primarily served as a technical means to drive overt and covert movement, rather than mandating a necessary correlation to rich overt morphology.
5 This is not entirely correct, however. Duarte (1995) shows that the inflectional paradigm in colloquial BP has undergone suppletion of the second and third persons during the 19th and 20th Centuries. However, the distinctions are reflected in writing.

2007; Grimshaw & Samek-Lodovici 1996; Holmberg 2010b; Samek-Lodovici 1996). A covert third person subject pronoun expresses topic continuity, whereas an overt subject pronoun introduces a new topic. The following Italian example from Holmberg & Sheehan (2010: 91) illustrates the point:

(3) *Gianni₁ dice che* (**lui₁*) *vuole comprare una machina.*
    Gianni   says that   he   wants   buy   a car
    'Gianni says that (he) wants to buy a car.'
    (Italian, Holmberg & Sheehan 2010: 91)

The example shows that the intended coreference between the embedded subject and the matrix subject obtains only if the embedded subject is null. If an overt pronoun resides in the embedded subject position, coreference is not possible. An overt subject pronoun is possible if it introduces a new topic (see also, Belletti, Bennati & Sorace 2007; Carminati 2002). Hence, the overt/covert dichotomy provides a classical example of complementary distribution. The following example illustrates that an identical pattern occurs in South Saami as well:

(4) *Piere₁   jeehti* (**dihte₁*) *sæjhta bijle-m   åestedh.*
    Piere.NOM said   3SG   wants   car-ACC buy
    'Piere said that he wants to buy a car.'

In order for the subject of the embedded clause to be coindexed with the higher subject, it must be null. In contrast, the overt subject pronoun *dihte* obligatorily triggers obviation.

The situation in partial NSLs like Finnish and Brazilian Portuguese (BP) is different. As shown in the examples in (5), an embedded subject in partial NSLs may be dropped when referring to the higher subject. However, subject drop is not enforced, and therefore an overt subject pronoun in the embedded clause may also be coreferential with the higher subject:

(5) a. *Juhani₁ kertoi että* (*hän₁*) *oli   ostanut omakotitalon.*
       Juhani said   that   he   PRF   bought house
       'Juhani said that he had bought a house.'
       (Finnish, Holmberg & Sheehan 2010: 131)
    b. *O João₁ disse que* (*ele₁*) *tinha comprado una casa.*
       the João said that   he   had   bought   a   house
       'João said that he had bought a house.'
       (Brazilian Portuguese, Holmberg & Sheehan 2010: 131)

Holmberg (2010b) develops a theory of consistent NSLs and partial NSLs that incorporates important ingredients from the works of Frascarelli (2007) and Roberts (2010) as well as ingredients from Holmberg (2005). We turn to these matters now.

2.2 Consistent NSLs

Following Frascarelli (2007), Holmberg (2010b) assumes that consistent NSLs have at their disposal an Aboutness-shift topic operator (henceforth TOP), which resides in the C-domain of the clause. The idea behind the proposal is that TOP mediates the relation between the null subject and the topic DP, which serves as the ultimate antecedent to the null pronoun. A schematic illustration is given in (6). Borrowing Sigurðsson's (2004, 2011, 2014a, b) terminology, the null pronoun in (6) is edge-linked to TOP, and TOP is contextually controlled by DP. This is indicated by indices.

(6) $\ldots$ DP$_1\ldots[_{CP}\ldots$TOP$_1\ldots[\ldots pro_1\ldots]]\ldots$

The specifics of Holmberg's theory reside in the relation between the pronoun and TOP. Holmberg (2005, 2010b) argues that small *pro* is a φP rather than a DP, drawing on the traditions of Cardinaletti & Starke (1999) and Déchaine & Wiltschko (2002), among others.[6] As such, φP is inherently incapable of referring, which is a privilege reserved for DPs. However, as the label suggests, φP is equipped with interpretable φ-features.[7] In contrast, T in consistent NSLs is equipped with an uninterpretable referential feature, [*u*D], in addition to uninterpretable φ-features, and, as a matter of course, tense features. Since the features of φP are a subset of the features on T, Holmberg argues on the basis of Roberts (2010), that φP incorporates into T under Agree. Consequently, φP does not raise to SpecTP, which does not project (compare, Alexiadou & Anagnostopoulou 1998):

(7)   a. $[_{TP} T_{[uD, u\varphi, \ldots]} [_{vP} \varphi P_{[3SG \ldots]} \ldots]]$
      b. $[_{TP} T_{[uD, u\varphi:3SG \ldots]} [_{vP} \cancel{\varphi P_{[3SG \ldots]}} \ldots]]$

---

6 For more recent work, see, for instance, Patel-Grosz & Grosz (2017).
7 It also has a case feature (see Holmberg 2010b), although I will set aside the issue of case in this chapter, as it will not have an impact on the overall issues under consideration.

However, at the stage depicted in (7b), [uD] on T has not been valued, which endangers the derivation. Once Frascarelli's TOP has merged into the C-domain, T is edge-linked to TOP which results in the valuation of [uD] on T. TOP in turn is contextually controlled by an accessible topic in the linguistic discourse. In examples like the Italian (3) and the South Saami (4) above, contextual control takes place under c-command, though c-command is not a necessary condition, as witnessed by (8) and (9):

(8)   *Gianni₁   non   ha   detto   niente,   ma   Paolo₂   ha   detto   che*
      Gianni₁   NEG   has   said   anything   but   Paolo   has   said   that
      ___₁/₂   *vuole   comprare   una   macchina   nuova.*
      (he)   wants   to.buy   a   car   new
      'Gianni₁ didn't say anything, but Paolo₂ said that he₁/₂ wants to buy a new car.'
      (Italian, Holmberg 2010b: 92)

The null subject in the embedded clause in the right conjunct may successfully refer to *Gianni,* in addition to the c-commanding antecedent *Paolo.* Exactly the same pattern occurs in South Saami:

(9)   *Ij   Næjla₁   maamakt   jeahteme,   bene   Pöövle₂   jeehti*   ___₁/₂
      NEG.3SG   Nils   anything   said   but   Paul   said   (3SG)
      sæjhta   orre   bijle-m   åestedh.
      wants   new   car-ACC   buy
      'Nils₁ hasn't said anything, but Paul₂ says that he₁/₂ wants to buy a new car.'

That is, the embedded null subject may refer to either the non c-commanding antecedent *Næjla,* or to the c-commanding antecedent *Pöövle.* The important point, of course, is the fact that both *Næjla* and *Pöövle* are topics. The implication is thus that TOP is sensitive to topichood, rather than the narrowly syntactic notion of c-command. In fact, as suggested by Frascarelli (2007), TOP may, at least in a non-technical sense, be viewed as a null copy of an established topic.

     As mentioned previously, overt pronouns, such as South Saami *dihte* '(s) he' are assumed to be DPs. DPs are structurally more complex than φPs, and they have the ability to refer to an entity in their own right. Moreover, overt pronouns carry information that renders them subject to vocabulary insertion at PF (Embick & Noyer 2001; Halle & Marantz 1993). It is therefore clear that the featural and structural bundle that makes up a pronoun like *dihte* '3SG.NOM' does not constitute a subset of the features on T, and therefore incorporation is blocked. Rather, as a result of being probed by T, DP raises to SpecTP, which in addition to the valuation of [uφ] on T, also amounts to the valuation of [uD]. Since *dihte*

'(s)he' introduces a new topic, it subsequently raises into the C-domain, either overtly or covertly:

(10) $[_{CP}$ <dihte> $[_{TP}$ dihte T$_{[\text{uD, uφ, ...}]}$ $[_{vP}$ ~~dihte~~$_{[3SG,NOM...]}$ . . .]]

To recapitulate, the null subject *pro*, a φP, incorporates into T under Agree. T hosts a [*u*D] feature which is valued by an Aboutness-shift topic operator, TOP, located in the C-domain. TOP in turn finds an antecedent in a topic in the linguistic discourse. Hence, the null subject is at the tail of a topic chain. An overt pronoun, on the other hand, introduces a new Aboutness-shift topic. A further example that illustrates the consistent NSL property in South Saami is given in (11).[8]

(11)   a. $[_{DP}$ *Gåeskie gonnoeh maakh*]$_1$ *lægan*   *aadtjen* $[_{DP}$ *ohtje niejtetje-m*]$_2$
             aunt and uncle   have.2DU   recently   little baby.girl-ACC
             åådtjeme.[9]
             gotten
             'My aunt and uncle just had a little baby girl.'
       b.   *Dihte*$_2$ / *___$_2$ *daan*   *minngemes*   *aejlegen*   *kristesovvi.*
             3SG / (3SG)   this   last   Sunday.GEN   was baptized
             'She was baptized last Sunday.'
       c.   *Jih*   *dellie*   <TOP$_2$>   ___$_2$ / *dihte*$_2$ *nomme-m*   *åadtjoeji.*
             and   then   (3SG) / 3SG   name-ACC   received
             'And then she received a name.'

In (11a), we find a complex subject DP, a topic. This DP triggers dual agreement on the finite verb. We also find the object *onne niejtetjem* 'little baby girl'. The pronominal subject in (11b) refers to the object in (11a), hence introducing a new topic. Therefore, the overt pronoun *dihte* is chosen over the null subject. In (11c), the subject refers to the topic introduced in (11b), and consequently the null subject is given priority.

---

8 Holm-Bull, Ella. 1986. Åarjelsaemien mubpie gærja. Saemien ööhpehtimmieraerie: Gærhkoe-jih ööhpehtimmie-departemente. [Holm-Bull, Ella. 1986. South Saami the second book. The Saami Council of Education: the Ministry of the Church and Education].

9 Notice that the conjunction *gonnoeh* (also *gon*) 'and' is exclusively used to conjoin nouns that denote kinship. Number and Case are signalled on the right-most head in the complex NP. Nouns can also be conjoined by a zero-conjunct, and the commitative case may also serve as a conjunct. The default conjunct is *jih* 'and' it is used in the same manner as in English. All the members of a complex *jih*-NP carry number and case inflection.

2.3 Partial NSLs

Partial NSLs are different. The fact that null subjects typically cannot be dropped in matrix clauses suggests that Frascarelli's TOP is not implicated:

(12) *(Hän) puhuu   englantia.
      (s/he)   speaks   English
      'S/he speaks English.'
      (Finnish, Holmberg 2005: 539)

Holmberg argues that Finnish has a null pronoun φP, but at the same time, T lacks [uD]. As a result, third person referential *pro*-drop of the consistent NSL variety is ruled out; instead, φP occurs in non-referential generic contexts.[10] Referential null subjects typically occur in embedded contexts, which was shown in (5) above. In such configurations, a null subject can alternate with an overt subject pronoun, without necessarily affecting coreference. The overt third person pronoun is argued to be a full DP, of the same kind as in consistent NSLs. In addition, Holmberg (2010b) and Holmberg & Sheehan (2010) assume that partial NSLs also have at their disposal a null [uD] pronoun, whose distribution essentially is restricted to embedded contexts. The [uD] feature on this null pronoun is assumed to be valued by the antecedent DP in the higher clause, as schematically shown in (13).

(13) $DP_1 \ldots [_{CP} [_{DP} u\text{D}]_1 [_{TP} <[_{DP} u\text{D}]> T [_{vP} <[_{DP} u\text{D}]> \ldots]]$

Consequently, antecedent linking in (13) is technically different from the instance of contextual control in (6) above. Holmberg (2010b) and Holmberg & Sheehan (2010) label null subjects in partial NSLs as control, but they are careful to clarify that the phenomenon is distinct from de facto control in a number of aspects.

2.4 Summary

This section presented the basic workings of the theory of null subjects developed in a series of works by Anders Holmberg. It was also shown that South Saami is a consistent NSL, on a par with, for instance, Italian.

---

10 I will not discuss generic and impersonal pronouns in this chapter.

3 Subject pronouns in complements of directive and commissive verbs

This section introduces the essential facts concerning the distribution of subject pronouns in finite complements of directive and commissive verbs in South Saami.

3.1 Directive verbs

As we noted in the introductory section, South Saami, a consistent NSL, has a variety of null subjects that are sharply distinct from ordinary third person referential null subjects in the language. The following example illustrates the issue:

(14) *Tjidtje Piere-m*₁ *stilli* ___₁ *edtji* *bijle-m dåvvodh.*
mother Piere-ACC ordered (3SG.M) WOLL.PST.3SG car-ACC repair
'Mother ordered Piere to repair the car.'

In (14), the null subject in the embedded clause is allowed to be coreferential with the object in the matrix clause. Prima facie, this is contrary to what is expected from a consistent NSL, since the object DP does not qualify as a topic. Further examples that illustrate the same points are given (15):

(15) a. *Læjsa Piere-m*₁ *mujhtiehti* _____₁ *edtji* *mielhke-m*
Lisa Piere-ACC reminded (3SG) WOLL.PST.3SG milk-ACC
åestedh.
buy
'Lisa reminded Piere to buy milk.'
   b. *Pöövle Maarja-m*₁ *soptsestehti* _____₁ *edtji*
Paul Mary-ACC persuaded (3SG) WOLL.PST.3SG
plaerie-m lohkedh.
newspaper-ACC read
'Paul persuaded Mary to read the newspaper.'

The cases at hand bear a striking resemblance to what Landau (2004) dubs as finite control. Although I will show that control per se is not involved in (14) and (15), i.e. the embedded subject is not PRO, it is noteworthy that each sentence has a directive reading. In fact, the matrix predicates (14) and (15) are bona fide control verbs that also license standard control infinitives:

(16) a. *Tjidtje   Piere-m₁   stilli   PRO₁ bijle-m   dåvvodh.*
      Mother   Piere-ACC ordered   car-ACC   repair
      'Mother ordered Piere to repair the car.'

   b. *Læjsa₁   Piere-m₁   mujhtiehti PRO₁ mielhke-m   åestedh.*
      Lisa   Piere-ACC reminded   milk-ACC   buy
      'Lisa reminded Piere to buy milk.'

   c. *Pöövle Maarja-m₁ soptsestehti PRO₁ plaerie-m   lohkedh.*
      Paul   Mary-ACC   persuaded   newspaper-ACC read
      'Paul persuaded Mary to read the newspaper.'

Landau (2004) points out that a necessary condition on finite control in Hebrew
is the presence of the future tense in the complement clause, as shown by the
following contrast:

(17) a. *Gil   hizkir   le-Rina₁   še-   ___₁/*₂   tin'al   et   ha-delet.*
      Gil   reminded to-Rina   that (3SG.F) will.lock.3SG.F   ACC   the-door
      'Gil reminded Rina to lock the door.'

   b. *Gil   hizkir   le-Rina₁   še-   *(hi₁/₂) na'ala   et   ha-delet.*
      Gil   reminded to-Rina   that 3SG.F   locked.3SG.F   ACC   the-door
      'Gil reminded Rina that she had locked the door.'
      (Hebrew, Landau 2004: 816–817)

In (17a), where the Control reading obtains, the embedded clause is in the future
tense. In contrast, the absence of the future tense in (17b) results in a non-direc-
tive, communicative reading. It is also noteworthy that the subject of the comple-
ment clause in (17b) cannot be null. A parallel pattern occurs in South Saami. The
observant reader will have noticed that the examples in (14) and (15) all involve
the future auxiliary *edtjedh*. Consider (18), where the embedded clause is in the
pluperfect:

(18) *Læjsa₁ Piere-m₂   mujhtiehti   ___₁/*₂/dihte*₁/₂/Maarja lij*
      Lisa   Piere-ACC reminded   (3SG)/3SG /Mary   had.3SG
      mielhke-m   åasteme.
      milk-ACC   bought
      'Lisa reminded Piere that (s)he/Mary had bought milk.'

In (18), the null subject in the complement clause cannot refer to the matrix
object. Instead, it occurs at the tail of a topic chain whose ultimate head is the
matrix subject. If the embedded subject is overt, it is obligatorily obviative with
respect to the matrix subject, and as a result it introduces a new topic. In other

words, (18) exhibits the trademark properties of consistent NSLs. In short, (18) is incompatible with the directive readings observed in (14) and (15). Thus, South Saami is like Hebrew in restricting the Control-like directive interpretation to future tense complements.

   While there are some striking similarities between finite directives in Hebrew and South Saami, there are also differences. The contrast between the Hebrew examples cited from Landau (2004) in (17) above shows that the subject of the directive must be null (see (17a)), whereas when the plain communicative reading obtains (see (17b)), the embedded third person subject must be overt. In contrast, the South Saami directive tolerates the presence an overt subject pronoun:

(19) *Læjsa   Piere-m₁   mujhtiehti edtji   dihte₁/*₂/*Maarja*
      Lisa   Piere-ACC   reminded   WOLL.PST.3SG   3SG / Mary
      mielhke-m   åestedh.
      milk-ACC   buy
      (i)   'Lisa reminded Piere to buy milk.'
      (ii)   *'Lisa reminded Piere that 3SG / Mary would buy milk.'

There are two noteworthy details to pay attention to in (19). Firstly, the subject pronoun *dihte* '(s)he' obligatorily occurs to the right of the future auxiliary. This is quite remarkable, since the language does not enforce subject-verb inversion in any other syntactic context. The second detail is the fact that the subject pronoun *dihte* is obligatorily coreferential with the matrix object, and the occurrence of an R-expression such as *Maarja* is altogether impossible. Hebrew, in contrast, does not permit an embedded overt third person pronoun to refer to the matrix object:

(20) *Himlacti   le-Rina₁   še-   Dani₂/hu₂   yearšem₁/*₂*
      I.recommended   to-Rina   that Dani/3SG.M   will.register.3SG.M
      la-xug   le-balšanut.
      to.the-department   the-linguistics
      'I recommended to Gil that Dani/he should register to the linguistics department.'
      (Hebrew, Landau 2004: 813)

As illustrated in (20), the occurrence of a lexical third person subject brings about control suspension in Hebrew. Returning to South Saami, it is fully possible for a third person pronoun or an R-expression to occur in a future complement clause, however with the proviso that it occurs in the canonical subject position to the left of the auxiliary:

(21) *Læjsa   Piere-m*₁   *mujtiehti   dihte*₁/₂/*Maarja   edtji   mielhke-m*
　　　Lisa   Piere-ACC   reminded   3SG /Mary   WOLL.PST3SG   milk-ACC
åestedh.
buy
　　(i)　‘Lisa reminded Piere that 3SG /Mary would/was about to buy milk.’
　　(ii)　*‘Lisa reminded Piere to buy milk.’

When this happens, the Control-like property is suspended, and the communicative reading emerges. The facts presented thus far correctly predict that a sentence like (15a), repeated here as (22), is ambiguous:

(22) *Læjsa*₁ *Piere-m*₂ *mujhtiehti* ___₁/₂ *edtji*   *mielhke-m   åestedh.*
　　　Lisa   Piere-ACC   reminded   (3SG)   WOLL.PST.3SG   milk-ACC   buy
　　(i)　‘Lisa reminded Piere to buy milk.’
　　(ii)　‘Lisa reminded Piere that she would/was about to buy milk.’

The directive reading, (22)(i), where the null subject is anteceded by the matrix object *Piere*, corresponds to example (19), where the overt subject pronoun occurs in the post-Aux position. The communicative reading, (22)(ii), where the null subject refers to *Lisa*, is comparable to the case where the overt pronoun occurs in the pre-Aux position (see (21)), in that both manifest the default referential patterns of overt and covert pronouns.

　　To take stock, we have seen that the finite complement of a directive verb in a Control-like context has the following characteristics: (a) it requires the future auxiliary, (b) it allows both a covert and an overt third person pronominal subject to be co-referent with the addressee argument in the higher clause, and (c) when an overt subject pronoun occurs in the complement clause, overt subject-aux inversion applies.

3.2 Commissive verbs

So far the focus of the discussion has been on the complements of directive verbs. Let us now draw attention to the complements of commissive verbs with the goal to show that these exhibit the same patterns as complements to directive verbs, albeit with some minor differences.

　　Consider the example in (23), where we find the commissive verb *dåajvoe-htidh* ‘promise’, which occurs in the intermediate clause. Notice also the null subject in the complement of the commissive verb:

(23) *Næjla*₁ *veanhta Pöövle*₂ *dåajvoehti* ___₁/₂ *edtji* *moerh*
    Nils   thinks   Paul   promised   (3SG)   WOLL.PST.3SG firewood
    risjnjedh.
    gather
    (i)   'Nils thinks that Paul promised that he would gather firewood.'
    (ii)   'Nils thinks that Paul promised to gather firewood.'

This sentence exhibits the same kind of ambiguity that we have noticed with directive verbs. Reading (23)(i) is reminiscent of control suspension, such that the null subject may refer either to the subject of the highest clause, or the subject of the intermediate clause. The sentence may also have a Control-like interpretation, (23)(ii), where the null subject refers to the subject of the intermediate clause, similarly to what is observed in the control infinitive in (24).

(24) *Næjla*₁ *veanhta Pöövle*₂ *dåajvoehti* PRO*₁/₂ *moerh* *risjnjedh.*
    Nils   thinks   Paul   promised   firewood gather
    (i)   *'Nils thinks that Paul promised that he would gather firewood.'
    (ii)   'Nils thinks that Paul promised to gather firewood.'

Commissives also permit the occurrence of an overt subject in the complement clause, and consequently a sentence like (23) can be disambiguated by the word order criterion established previously. However, before we consider the details of word order, we must consider the issue of pronominal selection. Beginning with (25), where the overt subject precedes the future auxiliary, three pertinent facts can be observed.

(25) *Næjla*₁ *veanhta Pöövle*₂ *dåajvoehti* *satne*₁/₂/*₃/*dihte**₁/*₂/₃/*Jåvva*
    Nils   thinks   Paul   promised   LOG.3SG /3SG   /John
    edtji *moerh* *risjnjedh.*
    WOLL.PST.3SG firewood gather
    (i)   'Nils thinks that Paul promised that he/John would gather firewood.'
    (ii)   *'Nils thinks that Paul promised to gather firewood.'

Firstly, the ordinary overt third person pronoun *dihte* can occur in the present context. However, it cannot be co-construed with a subject, but it must refer deictically. Secondly, a full R-expression may also fill the subject position of the most deeply embedded clause. Thirdly, when co-construal with a subject is intended, the third person logophoric pronoun *satne* is required. The default usage of *satne* is fully comparable to that of West African style logophoric pronouns (Vinka 2018, 2019). Under logophoric licensing, in the sense of Koopman & Sportiche (1989),

Adesola (2005), Baker (2018), among several others, *satne* refers to a c-command-ing, but not necessarily a local subject of an attitude predicate whose complement contains the logophoric pronoun.[11] Thus, *satne* in (25) may refer to either the subject of the intermediate clause, or the subject of the highest clause. One crucial difference between a referential null subject and a logophoric subject resides in whether or not the antecedent must c-command the subject pronoun. The following example shows that the logophoric *satne* cannot find its antecedent in the subject of the leftmost conjunct, since *satne* is not contained in the complement of *jeahteme* 'has said', (26).

(26) *Ij*   *Næjla$_1$*   *maam-akt*   *jeahteme,*   *bene*   *Laara$_2$*   *veahnta*
      NEG.3SG   Nils   anything   has.said   but   Lars   thinks
      Pöövle$_3$ *dååjvoehti*   ___$_{1/2/3}$/*satne*$_{*1/2/3}$/*dihte*$_{*1/*2/*3}$/*Jåvva*   *edtji*
      Paul   promised   (3SG)/LOG.3SG /3SG   /John   WOLL.PST.3SG
      moerh   *risjnjedh.*
      firewood   gather
      'Nils hasn't said anything, but Lars thinks that Paul promised that he/John would / was about to gather firewood.'

In short, (25) and (26) are representative examples of the default strategies of reference tracking in South Saami.

   Let us now turn to the second word order available, namely the one where the overt subject of the complement clause of the commissive verb occurs to the right of the future auxiliary:

(27) *Næjla$_1$*   *veanhta*   *Pöövle$_2$*   *dååjvoehti*   *edtji*
      Nils   thinks   Paul   promised   WOLL.PST.3SG
      satne$_{*1/2}$/*dihte*$_{1/2}$/*Jåvva*   *moerh*   *risjnjedh.*
      LOG.3SG/3SG /John   firewood   gather
      (i)   *'Nils$_1$ thinks that Paul promised that he$_1$ would gather firewood.'
      (ii)   'Nils thinks that Paul promised to gather firewood.'

We immediately notice that in (27), the only overt subject that is tolerated in the post-auxiliary position is the third person logophoric pronoun *satne*. Furthermore, the possible antecedent relations in (27) are severely constrained in com-

---

11 It should be noted that many languages with logophoric systems, such as Yoruba, allow logophoric licensing in the context of connected discourse (Safir 2004; Adesola 2005), in which a logophoric operator takes scope over the discourse. I set this issue aside here.

parison to (25) and (26). In spite of the logophoric properties of *satne* in (25) and (26) above, the default logophoric pattern is suspended in (27), such that *satne* can only refer to the subject in the next clause up. Indeed, the only reading available is the Control-like interpretation (27)(ii).

To summarize, when a directive or commissive verb combines with a finite complement clause in the future, a Control-like interpretation is available. If the complement clause has a null subject, cases of structural ambiguity may arise. However, if an embedded overt subject pronoun occurs to the right of the future auxiliary, only the Control-like interpretation is available. We have also noticed that the choice of overt subject pronouns is contingent on whether the higher verb is directive or commissive.

4 The future auxiliary as a subjunctive exponent

One peculiar characteristic of the Control-like finite complements under investigation is the fact that they require the presence of the future auxiliary *edtjedh* 'WOLL.INF'. As Landau (2004) notes, a necessary condition on finite control in Hebrew is that the complement clause is in the future tense. Landau continues by arguing that the future tense in Hebrew is the exponent of the subjunctive mood. By parity of reason, I claim that the South Saami future auxiliary, too, expresses the subjunctive mood.

4.1 Subjunctivity and the interpretation of tense

The subjunctive is usually described as an irrealis mood, in contrast to the indicative. Furthermore, the subjunctive typically occurs in embedded contexts, and its temporal interpretation is either constrained by or anaphoric on the tense of the higher clause. Many languages have designated subjunctive inflectional paradigms, for instance Italian and Spanish. Other languages may express the subjunctive by other means, such as a particle-like element, as in the Balkan languages, or by a designated complementizer, as in Russian (Shushurin 2017). Landau (2004), in turn, argues that the future tense in Hebrew instantiates the subjunctive in the complement of directive and commissive verbs, noticing that these constellations exhibit a number of properties that are at the core of subjunctivity, notwithstanding an otherwise fairly broad range of heterogeneity. For instance, subjunctives are typically temporally constrained, and exhibit sequence of tense

effects (SoT).[12] On a related note, Stowell (1982: 562) observes that the infinitival complement of a control verb signals a possible future, relative to the tense of the higher verb (see also Bresnan 1972). Indeed, various scholars have argued that one salient property of the subjunctive is that it is forward shifting, which at least in certain scenarios gives rise to irrealis readings (Kempchinsky 2009). The South Saami complements under consideration fit well into this picture. Consider the following example, where the main clause is in the past tense, and the continuations in the (a) and (b) examples are in the past and present respectively:

(28) *Tjidtje   Pöövle-m₁   stilli. . .*
     mother   Paul-ACC   ordered
     a. ___₁   *edtji*   *bijle-m   dåvvodh.*
         (3SG)   WOLL.PST.3SG   car-ACC   repair
         'Mother ordered Paul to repair the car.'
     b. ___₁   *edtja*   *bijle-m   dåvvodh.*
         (3SG)   WOLL.PRS.3SG   car-ACC   repair
         'Mother orders Paul to repair the car.'

The forward shifting character of the auxiliary *edtj-* 'woll' in both (28a) and (28b) signals that the subjunctive complement is posterior in relation to the past tense of the event in the higher clause. That is, mother's issuing of the order occurs prior to the car repairing event. I assume along the lines of Abusch (1997) and Wurmbrand (2014), among others, that the future auxiliary is composed of two parts. One component is the actual tenses [present] and [past]. The other component is the future modal itself, *edtj-*, which corresponds to Absusch's abstract *woll*. First, consider the ambiguous (28a), where *edtj-* qua *woll* has combined with [past].[13] On the one hand, it can convey that the car repairing event has been carried out by Paul at the utterance time. It is also compatible with a reading where the repairing event has not been carried out at the time of the utterance. Under the latter reading, the embedded past tense is semantically vacuous.[14] For

---

12 The SoT phenomenon has received much attention in contemporary generative grammar, for instance Abusch (1997), Demirdache & Uribe-Etxebarria (2005), Higginbotham (2001), Giorgi (2009), Kempchinsky (2009), Zagona (2014), among others, as well as in pre-generative eras, such as Reichenbach (1947). Due to the scope of inquiry of this chapter, I cannot do full justice to the rich body of research on the topic.
13 Under the line of analysis pursued in this chapter, the complement in (28a) is a past subjunctive, rather than an instance of future-in-the-past. As I will make clear, the embedded past tense in (28a) is not necessarily interpreted, a matter contingent on the possible application of LF deletion.
14 Subjunctive complements in South Saami do not readily allow temporal adverbial modification. Although I will not provide an account for this fact, a tentative conjecture is that it might

the present purposes, I follow Wurmbrand (2014) (who adopts the account developed in a series of works by Ogihara, e.g. Ogihara 2007), and assume that the SoT effect arises when the embedded tense is deleted at LF.[15] The idea is that a tense can delete when it is within the scope of another, identical tense. If the embedded tense has undergone LF deletion, the car repairing event in (28a) will have an irrealis interpretation at the utterance time. The embedded tense also has the option not to delete. In this case, the car repairing event will be understood as completed at time of the utterance. The present future complement clause in (28b), however, only conveys the meaning that the event of the complement clause has not been carried out. This is expected under the Ogihara-Wurmbrand approach. Since the embedded tense is distinct from the matrix tense, LF-deletion cannot apply. As a consequence, the embedded car repairing event cannot be interpreted as completed. This take on the future has the important implication that the future has an independent realization, which is distinct from the present and past tenses. In short, both (28a) and (28b) are perfectly well-formed directives, but they do differ in fine details of interpretation. Importantly, SoT is not strictly enforced in (28). Anticipating Section 5.3, the fact that SoT can be violated in South Saami subjunctive complement CPs suggests that the embedded tense is dependent, rather than anaphoric (Landau 2004).

Turning to (29), the situation is somewhat different. Here, the matrix clause is in the present tense, whereas the complement in (29a) is in the past tense, and in (29b) it is in the present tense.

(29) *Tjidtje₁ Piere-m₂ stillie. . .*
      mother   Piere-ACC   orders
   a.   *___₂ *edtji*   *bijle-m*   *dåvvodh.*
        (3SG)   WOLL.PST.3SG   car-ACC   repair
        'Mother orders Piere to repair the car.'
   b.   ___₂ *edtja*   *bijle-m*   *dåvvodh.*
        (3SG)   WOLL.PRS.3SG   car-ACC   repair
        'Mother orders Piere to repair the car.'

As we can see, (29a) is not a licit continuation of the main clause, since the embedded tense is [past]. Since the embedded tense is distinct from the matrix tense, LF deletion cannot apply and the sentence cannot be salvaged. In contrast,

---

follow from the analysis given in Section 5, where I argue that subjunctive complement clauses in the language are truncated, rather than being full CPs. To settle the issue, further research is required.

15 For a critique of the LF deletion account, see Zagona (2014).

if the embedded clause is [present], (29b), the forward shifting property is upheld and the sentence is grammatical. Notice further that (29b) cannot mean that the car repairing event has been carried out.

The same patterns of temporal interpretation are repeated in commissives, as is shown in (30), where the commissive verb is in the past tense, and (31) where it is the present tense.

(30) *Pöövle*₁   *dåajvoehti* . . .
    Paul   promised
    a.   ___₁   *edtji*   *bijle-m*   *dåvvodh.*
        (3SG.M)   WOLL.PST.3SG   car-ACC   repair
        'Paul promised to repair the car.'
    b.   ___₁   *edtja*   *bijle-m*   *dåvvodh.*
        (3SG.M)   WOLL.PRS.3SG   car-ACC   repair
        'Paul promised to repair the car.'

(31) *Pöövle*₁   *dåajvohte* . . .
    Paul   promises
    a.   *___₁   *edtji*   *bijle-m*   *dåvvodh.*
        (3SG.M)   WOLL.PST.3SG   car-ACC   repair
        'Paul promises to repair the car.'
    b.   ___₁   *edtja*   *bijle-m*   *dåvvodh.*
        (3SG.M)   WOLL.PRS.3SG   car-ACC   repair
        'Paul promises to repair the car.'

Again, we see that when the higher verb is in the past tense, the future auxiliary may be either in the past, (30a), or in the present (30b). Also, if the higher verb is in the present tense, the future auxiliary in the complement clause too must be in the present tense, (31a) vs (31b). The interpretational options are identical to (28) and (29). That is, SOT may be triggered if the embedded tense is identical to the matrix tense.

In sum, in the spirit of Landau (2004), I have argued that the future component *edtj-* qua *woll* serves as the exponent of the subjunctive mood in the complements under consideration. The facts that pertain to the temporal interpretation of the complements of directive and commissive verbs are consistent with the subjunctive analysis, where the embedded tense is constrained by the matrix tense. The facts presented suggest that subjunctive complements in South Saami are dependently tensed, rather than anaphorically tensed, in the sense of Landau (2004). These issues will play an important role in Section 5.

4.2 Subjunctivity and word order

Having argued that the future auxiliary in the complements under consideration expresses subjunctivity, we shall now turn our attention to another consequence of this hypothesis, namely word order. I have shown above that, in addition to the fact that the future auxiliary is obligatory, the complement of directive and commissive verbs tolerates an overt subject pronoun, with the proviso that the auxiliary precedes the pronoun. This word order quirk correlates well with the reasonably broad consensus in the syntactic literature that subjunctivity is encoded in a head, M(ood), which is located above T (see among several others, Cinque 1999; Giorgi & Pianesi 2004; Krapova 2001; Terzi 1997), as illustrated in (32).

(32)   . . . V [. . . M . . . [$_{TP}$ . . . T . . . vP]] . . .

The general idea in these works is that the Mood and the Tense heads communicate or Agree, with the ensuing result that T under certain circumstances raises to M. As an illustration, consider the following Greek example:

(33) i   Maria elpizi na   ftasi   o Janis   stin   ora   tu
      the Maria hopes   SBJV   arrives the Janis in.the   time   his
      'Maria hopes that John will arrive in time.'
      (Greek, Spyropoulos 2007: 164)

In Greek, a lexical DP in a subjunctive clause may surface to the right the verb. Terzi (1992) assumes that the VS word order is derived via V-to-T raising, and subsequent raising of the V-T complex to the position of the subjunctive particle. Giorgi & Pianesi proposes a similar analysis for Italian subjunctives where complementizer deletion has taken place.

In this chapter, I will assume the gist of (32) for South Saami. That is, the future auxiliary in the complement clause in (34) has undergone raising from T across the subject pronoun, to a higher position corresponding to M in (32).

(34) *Aahka*   *Pöövle-se₁* *jeehti* [*edtji*   [$_{TP}$ *dihte₁* t$_{edtji}$
      grandmother   Paul-ILL   told.3SG WOLL.PST.3SG   3SG
      moerh   *risjnjedh*]. . .].
      firewood   gather
      'Grandmother told Paul to gather firewood.'

In other words, the hypothesis that South Saami directive and commissive verbs can select subjunctive complement clauses provides a straightforward account

for the obligatory occurrence of the future auxiliary, since it is this item that is the lexical exponent of subjunctivity. Furthermore, the subjunctivity hypothesis also sheds light on the availability of SOT effects and the fact that subject-aux inversion applies obligatorily in the complement clause.

4.3 Which verbs can select subjunctive complements?

Up until now, the focus has been on elucidating the basic descriptive facts of finite complements of directive/commissive verbs. As was mentioned earlier, verbs that can select a subjunctive complement can also select an infinitival complement. Hence, the subjunctives in (35a) vary with the control infinitives in (35b).

(35) a. *Tjidtje*   *Piere-m*$_1$ *stilli/soptsestehti/mujhtiehti/meedti* *edtji*
mother   Piere-ACC ordered/persuaded/reminded/ask   WOLL.PST.3SG
(*dihte*$_1$) *bijle-m*   *dåvvodh.*
3SG   car-ACC   repair
'Mother ordered/persuaded/reminded/asked Piere to repair the car.'
  b. *Tjidtje*   *Piere-m*$_1$ *stilli/soptsestehti/mujhtiehti/meedti*   PRO$_1$
mother   Piere-ACC ordered/persuaded/reminded/asked
bijle-m   *dåvvodh.*
car-ACC   repair
'Mother ordered/persuaded/reminded/asked Piere to repair the car.'

It is also important to note that subjunctive selection does not generalize to all object control verbs. Causative and permissive object control verbs such as *noerkedh* 'force' and *baajedh* 'allow' are incompatible with subjunctive complements, as in (36a), but only tolerate infinitival complements, (36b):

(36) a. *Tjidtje   *Piere-m*$_1$ *noerhki/beeji*   *edtji*   (*dihte*$_1$)
mother   Piere-ACC forced/allowed WOLL.PST.3SG 3SG
bijle-m   *dåvvodh.*
car-ACC   repair
'Mother forced/allowed Piere to repair the car.'
  b. *Tjidtje*   *Piere-m*$_1$ *noerhki/beeji* PRO$_1$ *bijle-m* *dåvvodh.*
mother   Piere-ACC forced/allowed   car-ACC   repair
'Mother forced/allowed Piere to repair the car.'

Likewise, commissive verbs can select both subjunctive and infinitival comple-ments, as shown in (37), whereas implicative and aspectual subject control verbs are incompatible with subjunctives, (38).

(37) a. *Pöövle$_1$ dåajvoehti/sjæjsjali edtji* (*satne$_1$*) *moerh*
Paul promised/decided WOLL.PST.3SG LOG.3SG firewood
risjnjedh.
gather
'Paul promised/decided to gather firewood.'

b. *Pöövle$_1$ dåajvoehti/sjæjsjali* PRO$_1$ *moerh risjnjedh.*
Paul promised/decided firewood gather
'Paul promised to gather firewood.'

(38) a. **Pöövle$_1$* buektiehti/eelki/maahta edtji* (*satne$_1$*)
Paul managed/began/can WOLL.PST.3SG LOG.3SG
moerh risjnjedh.
firewood gather
'Paul managed/began/can (to) gather firewood.'

b. *Pöövle$_1$ buektiehti/eelki/maahta* PRO$_1$ *moerh risjnjedh.*
Paul managed/began/can firewood gather
'Paul managed/began/can (to) gather firewood.'

The dichotomy between control verbs in South Saami that can select subjunctive complements and those that cannot, correlates nicely to Landau's (2000) typology of control in terms of partial and exhaustive control. The verbs that may select sub-junctive complements in (35) and (37) are verbs that can give rise to partial control. On the other hand, the verbs that are incompatible with subjunctives, (36) and (38), yield exhaustive control only. Simply put, in partial control, the reference of PRO is not required to be exhausted by the reference of the controller, whereas in exhaustive control, the reference of PRO must be exhausted by the reference of the controller. The distinction plays out most clearly with predicates whose subjects inherently denote plurality. For instance, South Saami has a plain transitive verb *råakedh* 'meet.TR', which requires a subject and an object, (39a) versus (39b):

(39) a. *Jååktan, Pöövle Maarja-m råaki tsåaka golme.*
yesterday Paul Mary-ACC met.TR.3SG clock three
'Yesterday, Paul met Mary at three o'clock.'

b. **Jååktan, Pöövle jih Maarja råaki-n tsåaka golme*
yesterday Paul and Mary met.TR-3PL clock three
'Yesterday, Paul and Mary met at three o'clock.'

In addition, there is an intransitive verb *råakedidh* 'meet.INTR', which requires a plural subject, (40a) versus (40b).

(40) a. *Jååktan,   Pöövle jih   Maarja råakedi-n   tsåaka   golme.*
          yesterday Paul   and   Mary   met.INTR-3PL clock   three
          'Yesterday, Paul and Mary met at three o'clock.'

 b. **Jååktan   Pöövle (Maarja-m) råakedi   tsåaka   golme.*
          yesterday Paul   (Mary-ACC) met.INTR.3SG clock   three
          'Yesterday, Paul met (Mary) at three o'clock.'

Next, the examples in (41) show that the verb *dåajvoehtidh* 'promise' in addition to exhaustive control, also licenses partial control:

(41) a. *Pöövle$_1$ dåajvoehti PRO$_1$ Maarja-m råakedh   tsåaka   golme.*
          Paul   promised   Mary-ACC meet.TR   clock   three
          'Paul promised to meet Mary at three o'clock.'

 b. *Pöövle$_1$ dåajvoehti PRO$_{1+}$ råakedidh   tsåaka   golme.*
          Paul   promised   meet.INTR   clock   three
          'Paul promised to meet at three o'clock.'

In (41a) the antecedent *Pöövle* exhausts the reference of PRO, that is, PRO is referentially identical to the antecedent. However, this is not the case in (41b), where the verb in the embedded clause is the intransitive *råakedidh* 'meet.INTR'. We know from the contrast between (40a) and (40b) above that this verb requires a plural subject. Nevertheless, (41b) is fully grammatical, although the reference of PRO is not exhausted by the singular antecedent *Pöövle*. Thus, (41b) is a case of partial control.

Turning to the Control verb *buktiehtidh* 'manage' in (42), we see that only exhaustive control is permitted:

(42) a. *Pöövle$_1$ buektiehti PRO$_1$ Maarja-m råakedh   tsåaka   golme.*
          Paul   managed   Mary-ACC meet.TR   clock   three
          'Paul managed to meet Mary at three o'clock.'

 b. **Pöövle$_1$ buektiehti PRO$_{1+}$ råakedidh   tsåaka   golme.*
          Paul   managed   meet.INTR   clock   three
          'Paul managed to meet at three o'clock.'

 c. *Dah$_1$ buektiehti-n PRO$_1$ råakedidh   tsåaka   golme.*
          they   managed-3PL   meet.INTR   clock   three
          'They managed to meet at three o'clock.'

In the grammatical (42a) the reference of PRO is identical to the antecedent *Pöövle*, and hence the criteria for exhaustive control are satisfied. In contrast, (42b) is ruled out. This is so because PRO by necessity denotes plurality by virtue of being the subject of *råakedidh* 'meet.INTR', whereas the antecedent is singular. In (42c) the antecedent is plural and, therefore, exhaustive control obtains.

The patterns we have identified with regards to subject control verbs in (41) and (42) carry over to object control verbs as well. To begin with, the directive verbs in (35) support both exhaustive and partial control, as shown in (43a) and (43b), respectively:

(43) a. *Tjidtje   Pöövle-m$_1$ stilli/soptsestehti/mujhtiehti/meedti   PRO$_1$*
mother   Paul-ACC   ordered/persuaded/reminded/asked
Maarja-m   råakedh   tsåaka   golme.
Mary-ACC   meet.TR   clock   three
'Mother ordered/persuaded/reminded/asked Paul to meet Mary at three o'clock.'
   b. *Tjidtje   Pöövle-m$_1$ stilli/soptsestehti/mujhtiehti/meedti   PRO$_{1+}$*
mother   Paul-ACC   ordered/persuaded/reminded/asked
råakedidh   tsåaka   golme.
meet.INTR   clock   three
'Mother ordered/persuaded/reminded/asked Paul to meet at three o'clock.'

In contrast, causative object control verbs such as *noerhkedh* 'force' and *baajedh* 'let' are only compatible with exhaustive control, as shown by the contrast between (44a) and (44b).

(44) a. *Tjidtje   Pöövle-m$_1$ noerhki/beeji   PRO$_1$ Maarja-m   råakedh*
mother   Paul-ACC   forced/allowed   Mary-ACC   meet.TR
tsåaka   golme.
clock   three
'Mother forced/allowed Piere to meet Mary at three o'clock.'
   b. **Tjidtje   Pöövle-m$_1$ noerhki/beeji   PRO$_{1+}$ råakedidh   tsåaka*
mother   Paul-ACC   forced/allowed   meet.INTR   clock
golme.
three
'Mother forced/allowed Paul to meet at three o'clock.'

c. *Tjidtje* [$_{DP}$ *Pöövle-m   jih   Maarja-m*]$_1$ *noerhki/beeji*   PRO$_1$
mother   Paul-ACC   and   Mary-ACC   forced/allowed
råakedidh   tsåaka   golme.
meet.INT   clock   three
'Mother forced / allowed Paul and Mary to meet at three o'clock.'

In particular, (44b) is out for the same reason as (42b). That is, under the hypothesis that *noerhkedh* 'force' and *baajedh* 'let, allow' license only exhaustive control, the reference of the plural PRO fails to be exhausted by the singular antecedent. However, if the antecedent is plural, then exhaustive control obtains, as shown in (44c) above.

In sum, control verbs that license partial control infinitives are identical to the class of verbs that have the option of selecting subjunctive complements in South Saami. According to Landau (2004), this fact is not a coincidence, rather, an expected consequence of the distinction between partial and exhaustive control complements. Landau (2000) argues at length that the dichotomy is linked to tense. More specifically, partial control infinitives are abstractly tensed, whereas exhaustive control infinitives are untensed. Thus, verbs that take a partial control complement select, among other things, tense. Consequently, such verbs may combine either with a tensed infinitive or a subjunctive complement in South Saami.

4.4 Section summary

This section argued that the obligatorily occurring future auxiliary in subjunctive complements in South Saami is the lexical exponent of subjunctivity. This in turn has consequences for complement selection. It was shown that subjunctive complements in South Saami are selected by verbs that also license partial control infinitives, in the sense of Landau (2000).

5 Control-like subjunctives in South Saami do not involve control

In this section, I will present the analysis of the phenomena under consideration. I begin by considering Landau's (2004) theory of finite control, and I conclude that control is not at stake in South Saami subjunctives. The presentation continues by claiming that the complement clause is smaller than a CP, which is a

factor in accounting for the referential properties of the subject of a South Saami subjunctive complement, namely the fact that it must refer to an argument in the next clause up. Another claim is that the null subject in these contexts is not *pro*, but rather a PF-deleted overt subject pronoun.

5.1 Landau (2004) on finite control

In his seminal work on finite control, Landau (2004) distinguishes between two types of subjunctive complements. One variety carries dependent tense, which is a *de facto* tensed complement, albeit constrained by the temporal properties of the matrix clause. The other variety consists of an anaphorically tensed subjunctive complement. In contrast to the former, the latter is tenseless. Whether or not a subjunctive complement carries dependent or anaphoric tense depends on the properties of the selecting predicate. Dependently tensed subjunctives are selected by verbs that license partial control, whereas anaphorically tensed subjunctives are selected by verbs that license exhaustive control. Thus, there is a direct correlation between subjunctives and obligatory control. Landau argues further that partial control verbs select for an uninterpretable [+T] feature on C, which in the case of subjunctives results in dependent tense. On the other hand, exhaustive control verbs select for an uninterpretable [-T] on C, which yields anaphoric tense. The tense features on C and I enter into checking relation, where [±T] on I is interpretable. Landau's structure is illustrated in (45).

(45)   V. . . [$_{CP}$ C$_{[±T]}$ [$_{IP}$ I$_{[±T]}$ VP]]

Landau furthermore assumes that overt agreement is contingent on the presence [+T]. If both C and I are [+T] and I is [+Agr], then C too is [+Agr]. All things being equal, since the South Saami subjunctive is dependently tensed and since the future auxiliary carries overt agreement, it follows that both C and I are [+T, +Agr]. Landau also assumes a redundancy rule that assigns a referential feature, [±R], to the heads in question. The specification [-Agr], or its absence altogether, results in [-R]. In the case of South Saami, the presence of [+Agr] on C and I respectively renders these heads [+R]. [+R] on C and I is checked against a referential DP, i.e. a lexical DP or *pro*. Hence, the relevant details of South Saami in Landau's (2004) theory are as shown in (46), which is identical to Landau's (2004: 844) example (41a).

(46) $[_{CP} \ldots DP \ldots F \ldots [_{CP} C_{[+T, +Agr, +R]} [_{IP} I_{[+T, +Agr, +R]} [_{VP} DP/pro \ldots]]]]$

|_____| |_____| |_____|
Agree   Agree$_{[+T, +Agr, +R]}$ Agree$_{[+Agr, +R]}$

Landau's calculus correctly rules out PRO in South Saami subjunctive comple-
ments, and it also correctly predicts that the embedded subject could be either a
null or an overt pronoun. This follows from the fact that that C and I are [+T, +Agr],
and thus [+R], with the ensuing result that I Agrees with a [+R] *pro* or an overt
pronoun. Moreover, nothing in Landau's account bans coreference between the
DP in the higher clause and *pro*. However, with regard to South Saami, (46) fails to
account for the fact that the embedded pronominal subject is obligatorily corefer-
ential with an argument in the higher clause, in conjunction with the fact that full
lexical DPs qua R-expressions are illicit. Hence an alternative account is required.

5.2 Towards an analysis

It has become evident in the course of the presentation that South Saami has
three series of third person pronouns that interact in the overall system of refer-
ence-tracking (see Table 1). They are all gender neutral, and the overt members in
the series inflect for seven morphological cases and three numbers.

Table 1: 3rd person pronoun series interacting in reference-tracking.

Null pronouns	Logophoric pronouns	Obviative pronouns
3SG (s/he)	*satne*	*dihte*
3DU (they both)	*såtnoe*	*dah (guaktah)*
3PL (they)	*sijjieh*	*dah*

Null and obviative pronouns occur by default in indicative main clauses and com-
plements, whereas logophoric pronouns are restricted to indicative complement
clauses of attitude predicates. The most conspicuous trademark of subjunctive
complements in South Saami is the suspension of the standard mechanisms of
reference tracking, namely the consistent-NSL property and logophoricity (Vinka
2018). Third person referential *pro*-drop and logophoricity have in common that
their occurrence is contingent on the presence of a designated operator, or opera-
tor-like item, in the C-domain (e.g. Adesola 2005; Frascarelli 2007). Although the
two differ sharply in numerous aspects, they intersect distributionally in impor-
tant ways. For instance, they are restricted to non-subjunctive finite clauses, and
they allow non-local antecedents. These pronouns are recycled in subjunctive

complements in the language, where they behave differently from their defaults, but persistently refer to an argument in the next clause up. Assume for a moment that the null subject in a subjunctive essentially is derived in a manner that dovetails with the consistent NSL account, and as a result it has incorporated into T (Holmberg 2010b; Roberts 2010). The immediate issue that arises is how the uninterpretable feature on T is valued. It could be stipulated that there is an equivalent to Frascarelli's Aboutness-shift topic operator, which enters an Agree relation with T, with the effect that [uD] on T is valued.

(47)   [$_{CP}$ mother Paul$_1$-ACC ordered [$_{CP}$ OP$_{[D]1}$ FUT-T$_{[uD]}$-*pro* wash clothes]]

However, a challenge to the approach sketched in (48) is why the subject of the subjunctive complement clause of a directive verb freely oscillates between a null and an overt pronoun. What is particularly troublesome by appealing to a modification of the standard *pro*-drop account is that the third person pronoun *dihte* carries a referential feature and can refer, in the sense that it can introduce a new topic and trigger obviation. If one were to assume that the occurrence of an overt pronoun precludes the existence of the putative operator in (47), it is hard to explain why *dihte* is obligatorily co-referent with an argument in the next clause up. In short, there are good reasons to dismiss the idea that a modified version of consistent NSL-style *pro*-drop is involved in subjunctive complement clauses. In order to circumvent this problem, I propose that the null subject in subjunctive complements is derived by the deletion of an overt pronoun.

5.3 The subject in subjunctive complements

I will now outline the proposal that the null subject in a subjunctive complement is a PF-deleted overt pronoun. Three issues are of importance, namely the status of the subjunctive complement, the mechanism of reference tracking and the role of the future auxiliary.

Addressing the status of the subjunctive complement, I will follow the gist of Landau (2004) in assuming that the existence of tense in a dependently tensed subjunctive complement is a matter of selection, along the lines in (42) above. For the sake of argument, I follow Landau (2004) in assuming that the selected projection is bestowed with a feature, namely future [FUT].[16] I also depart from

---

[16] See Grano (2012) and Wurmbrand (2014) for criticism of Landau's treatment of tensed and untensed infinitives.

Landau (2004) and instead I assume that the subjunctive complement is a Mood Phrase (MP), along the lines of Giorgi & Pianesi (2004). This approach impacts on the second issue, namely the mechanism of reference tracking, since it has the effect of preventing the occurrence of Frascarelli's topic operator, which is taken to be housed in the C-domain. For the same reason, the subjunctive complement cannot host a logophoric operator, which too is an element of the C-domain (e.g. Adesola 2005; Anand 2006).

Landau (2015) recasts the distinction between partial and exhaustive control. He points out that verbs that select partial control complements are attitude predicates, and reversely, exhaustive control complements are selected by non-attitude predicates. In this theory, predicative and logophoric control substitute exhaustive and partial control, respectively. Simply put, predicative control involves a predication relation between an infinitival complement and a controller (see also Williams 1980). On the other hand, in logophoric control, the controller variable binds an abstract pronominal in the C-domain of the complement clause, which in turn enters a predication relation with an infinitival FinP. Landau (2015) argues further that the abstract pronominal, among other things, encodes the participant features [Author] and [Addressee]. The fact that attitude predicates can license elements of this kind has strong support in the literature on logophoricity and indexicality in general (Anand 2006; Baker 2018; Safir 2004; Schlenker 2003). However, my hypothesis that the subjunctive complement is an MP, rather than a CP, precludes the existence of an abstract pronominal of the kind mentioned. I will follow a proposal by Shushurin (2017), which preserves the spirit but not the details of Landau (2015), and posits that the participant feature is encoded on the head M, as shown in (48).

(48) $[_{CP} \ldots DP \ldots F \, [_{MP} M_{[FUT][AUTH/ADDR]} \, [_{TP} FUT\text{-}T_{[uD]} \, vP]]]$

With this background, I begin by considering the overt subject of a subjunctive complement of a directive verb. Following Holmberg (2010b) I assume that the overt third person pronoun *dihte,* by virtue of being a DP, agrees with the φ-features on T, and it values the referential feature on T. This is shown in (49) below. Notice also that raising into the specifier of TP takes place because the features of the pronoun are not a subset of the relevant features on T. Next, following Krapova (2001) and Giorgi & Pianesi (2004), among others, the T-complex complex raises into M, which is indicated in (49).

(49) $[_{CP} \ldots DP \ldots V_{Directive} \, [_{MP} FUT\text{-}T\text{-}M_{[ADDR]} \, [_{TP} \textit{dihte} \, t_{FUT\text{-}T} \, [vP \, t_{\textit{dihte}} \ldots]]]$

Furthermore, the MP is selected by the directive verb, and as a result the M head carries the feature [Addressee], which picks out the object addressee of the direc-

tive verb. This feature and the pronoun enters an Agree relation, which in Sig-urðsson's (2011) terms amounts to edge-linking. Consequently, the pronoun *dihte* will be coreferential with the object addressee argument in the higher clause, which contextually controls the participant feature.

There is one noteworthy detail with respect to pronominal choice. As I mentioned previously, the pronoun *dihte* triggers subject/topic obviation. Since *dihte* is not coreferential with a subject/topic in (49), the intended co-construal is possible. That is, the overt pronoun is compatible with the feature [ADDR(ESSEE)]. However, if the verb that selects the subjunctive complement is a commissive, then the participant feature on M is [AUTH(OR)]. In this case the logophoric pronoun *satne* occurs in the embedded clause:

(50) $[_{CP} \ldots DP \ldots V_{commissive} [_{MP} \text{F}\text{UT-T-M}_{[\text{AUTH}]} [_{TP} \text{satne } t_{\text{FUT-T}} [_{vP} t_{\text{satne}} \ldots]]]$

Logophoric pronouns are typically subject-oriented, such that they refer to the subject of an attitude predicate (Baker 2018; Diercks 2013).[17] Clearly, external arguments of this variety usually express the author of what is being said, thought, believed, etc. (Clements 1975:141). For the purposes of this chapter, it suffices to assume that the logophoric pronoun *satne* is equipped with an [AUTH(OR)] feature. The overt pronoun *dihte,* on the other hand, has no participant specification, and represents the elsewhere case. Thus, at PF, vocabulary insertion in the sense of Halle & Marantz (1993) applies as follows:

(51)   /satne/ ↔ [AUTH]
       /dihte/ ↔ elsewhere

I will now turn to the third issue, which tangibly addresses the role of the future auxiliary. As I showed in the previous sections, the general behavior of the alternation between overt and covert subject pronouns in South Saami subjunctive complements closely dovetails with the distribution of overt and covert subject pronouns in complement clauses in partial NSLs (Holmberg 2005, 2010b; Holmberg & Sheehan 2010),[18] as is illustrated in the Finnish and Brazilian Portuguese (BP) examples in (52a) and (52b), respectively.

---

17 There are languages where the addressee argument of an attitude predicate can antecede a logophoric pronoun (Andersen 1999; Culy 1994). However, this is not possible in South Saami.
18 Third person referential null subjects in partial NSLs have attracted much research over the past two decades or so. The literature on the topic is too rich to review within the scope of this chapter. However, in addition to the sources referred to above, see also Vainikka & Levy (1999), Modesto (2007), Brattico (2017) and Barbosa (2009, 2019), among several others.

(52) a. *Juhani₁ kertoi että (hän₁) oli ostanut omakotitalon.*
      Juhani said   that   he   had bought house
      'Juhani said that he had bought a house.'
      (Finnish, Holmberg & Sheehan 2010: 131)

    b. *O João₁ disse que (ele₁) tinha comprado uma casa.*
      the João said that he   had   bought a   house
      'João said that he had bought a house.'
      (Brazilian Portuguese, Holmberg & Sheehan 2010: 131)

In his study on argument drop, Sigurðsson (2011) pays particular attention to the fact that the presence of a lexicalized C is required in order for a null subject to be licit in partial NSLs. Sigurðsson (2011) argues that embedded null subjects in partial NSLs are of the same general ilk as dropped topics in Germanic. More specifically, Germanic argument drop is contingent on V2 and access to Spec,CP, by the Empty Left Edge Condition (Sigurðsson & Maling 2008). The primary difference between partial NSLs and Germanic is the fact that argument drop is essentially restricted to embedded domains in the former, whereas in the latter argument drop primarily applies in matrix contexts. A null subject in a partial NSL is contingent on the presence of an overt complementizer, as expected from Sigurðsson's lexicalized C condition. With this in mind, consider (53), which exemplifies this condition in the partial NSL North Saami (the examples below are from Partapuoli 2020).

(53) a. *Lisa₁ dajai ahte (son₁) sihtá oastit biilla.*
      Lisa said that she   wants to buy car
      'Lisa said that she wanted to buy a car.'

    b. *Lisa₁ dajai *(son₁) sihtá oastit biilla.*
      Lisa said she   wants to buy car
      'Lisa said that she wanted to buy a car.'
      (North Saami, Partapuoli 2020: 4)

Example (53a) is fully comparable to Finnish and Brazilian Portuguese in (52). However, if the complementizer is dropped, as in (53b), the embedded subject must be overt. Drawing on Sigurðsson (2011), the null subjects in (52) and (53a) are derived as follows. The pronoun moves into the C-domain to a position located to the immediate left of the overt complementizer. This movement is necessary in order for the pronoun to be accessible for control from the higher clause (Sigurðsson & Maling 2008: 295). The pronoun, in turn, cliticizes onto the complementizer and subsequently deletes at PF. The other scenario is illustrated in (53b), where C is non-lexical. Assume that the subject pronoun raises string-vacuously into the

C-domain, across the abstract the C-head. However, since C is null, cliticization of the pronoun fails, due to the lack of a lexical host. Therefore, the occurrence of a null subject pronoun is impossible in (53b) (see also Shushurin 2017).

When all facts are considered, a possible approach is that the null subject in a South Saami subjunctive complement is a PF-deleted overt subject pronoun, on a par with the non-subjunctives in (52) and (53a). One potential objection that could be raised is the fact that South Saami lacks a default complementizer corresponding to the English *that*. However, the syntax of subjunctive complements provides an answer as to how the lexicalized C condition of Sigurðsson (2011: 296) is met, even in the absence of C. The key lies in the fact that the future auxiliary in (49) and (50) above, raises to M. Hence, the head to the left of the subject pronoun is lexicalized at the relevant level. Assume now that the pronoun, whether it be *dihte* in the directive (49) or *satne* in the commissive (50), raises past the auxiliary and cliticizes onto it as shown in (54).

(54) $[_{CP} \ldots DP \ldots V_{Directive} [_{MP}$ ***dihte***-$F\text{UT}$-T-$M_{[ADDR]} [_{TP} t_{dihte} t_{FUT\text{-}T} [vP t_{dihte} \ldots]]]$

The pronoun subsequently deletes, as indicated by strike-throughs in (54).[19] Once the pronoun cliticizes to M, it must delete.[20]

6 Unselected subjunctives in South Saami

Before concluding this chapter, I will provide a brief comparison of selected and unselected subjunctives in South Saami. It is by now clear that selected subjunctives are restricted to the complement of directive and commissive verbs. However, the language also has access to unselected subjunctives, which uniformly serve as purpose clauses, as shown in (55).

---

19 My use of the notion PF-deletion may seem somewhat misleading. Since I assume a DM style model where vocabulary insertion by definition takes place at PF, PF deletion is better understood as failure to undergo vocabulary insertion. See Merchant (2001) for a related issue in the analysis of ellipsis phenomena.

20 The analysis is not without complications. For instance, there is no solid empirical foundation for positing pronominal cliticization in South Saami for the simple, but nonetheless important, reason that the language lacks pronominal clitics. I will however have to leave this issue for future research.

(55) (*Jupmele*) *voesse-m saemien*₁ *gåmmese vedti* [*edtji*
Creator bag-ACC Saami to woman gave WOLL.PST.3SG
⋆(*dihte*₁) *dam geehtedh*].
3SG it protect
'The Creator gave the Saami woman the bag, in order for her to protect it.'

Subjunctive purpose clauses have the same surface profile as subjunctive complement clauses. That is, they are introduced by the future auxiliary, which is immediately followed by the subject. However, unlike the situation in complement clauses, the subject pronoun in (55) cannot be null and retain reference to the goal argument in the matrix clause. If the subject is null, it must refer to the *Creator*. That is, the mechanism of reference tracking falls under the consistent-NSL module in subjunctive purpose clauses. Under the assumption that purpose clauses are adjuncts, it is reasonable to assume that they are full CPs. This in turn explains why topic-chained third person referential null subjects are possible:

(56) *Gåmma*₁ *roehtedi dan tjarke guktie åajsoeji* [*edtji*
woman ran as fast as could WOLL.PST.3SG
(⋆*dihte*₁) *tjoejhkide voessese vihth tjöönghkedh*].
3SG mosquitoes into bag again gather
'The woman ran as fast as she could, in order to gather the mosquitoes back into the bag again.'

The example in (56) shows that the null subject of the purpose clause refers to the subject/topic in the matrix clause. Hence, the patterns in (55) and (56) can receive a straightforward explanation if subjunctive purpose clauses are CPs that provide room for a Frascarelli-type Aboutness-shift topic operator that mediates the relation between the null subject and the antecedent *gåmma* 'woman' in (56). By parity of reasoning, the overt subject in (55) introduces a new topic.

Finally, subjunctive purpose clauses can be negated, consider (57):

(57) *Pöövle tjoeri maanide*₁ *geehtedh olles dah*₁
Paul had to the children keep.an.eye.on NEG.SBJV.3SG they
tjiektjh låavtegen sontere.
kick teepee.GEN against
'Paul had to keep an eye on the kids, in order for them not to kick at the teepee.'

The sentential negation in South Saami is high and it occurs at the left edge of the clause (Vinka 2007). The negation in subjunctive purposives is expressed by the clause-initial item *oll-* 'NEG.SBJV'. Since the subjunctive negation precludes the future auxiliary, I assume for now that the M head is bundled with the negation.[21] However, a subjunctive complement cannot be negated:

(58) * *Tjidtje   Pöövle-m$_1$ stilli   olles   (dihte$_1$)   tjiektjh*
      mother   Paul-ACC   ordered   NEG.SBJV.3SG   (3SG)   kick
      låavtegen   sontere.
      teepee.GEN   against
      'Mother ordered Paul not to kick at the teepee.'

The ill-formedness of (58) follows immediately from the assumption that selected subjunctives are not full CPs, but rather MPs. In contrast, if subjunctive purpose clauses are CPs, then the fact that the negation in (57) is licit, is accounted for.

7 Concluding remarks

In this chapter I claimed that there are two types of third person null subjects in South Saami. The null subject in indicative clauses as well as in unselected subjunctive clauses is a *pro* φP, in the sense of Holmberg (2010b), whereas the null subject in selected subjunctive complements is an overt pronoun that escapes pronunciation by cliticizing to the Mood head in a subjunctive clause. The argumentation has been along the following lines. Directive and commissive verbs may select subjunctive complements and the future auxiliary serves as the exponent of subjunctivity, in conjunction with subject-aux inversion. Moreover, finite subjunctive complements are smaller than CP, which has repercussions for the mechanisms of reference tracking, such that the default consistent-NSL property and logophoricity are suspended in the contexts investigated. The null subject in a subjunctive complement is a PF- deleted, or unpronounced, overt subject pronoun, which is facilitated by a participant feature [AUTH/ADDR] and cliticization onto the future auxiliary. The participant feature also plays an important role in accounting for the fact that the subject of the subjunctive complement is obligatorily coreferential with an argument in the higher clause.

---

21 The bundling effect can be conceived of in terms of raising, or spanning (Merchant 2015). Notice furthermore that the subjunctive negation is distinct from the indicative and imperative negations, *i-* and *aellie-* respectively.

The issues presented here also contribute to the puzzle and typology of subjunctives in general, which are notoriously diverse. Another ambition has been to contribute to the description of South Saami. Notwithstanding that the language is both understudied and endangered, its combination of third person referential *pro*-drop and West African style logophoricity in conjunction with the referential patterns in subjunctive complements, provide important probes into reference tracking in general.

Needless to say, there are several issues that remain theoretically unresolved and require further attention.

References

Abusch, Dorit. 1997. Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20. 1–50.

Adesola, Oluseye. 2005. Pronouns and null operators. PhD dissertation. Rutgers University, New Brunswick, NJ.

Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Word order, V-movement and EPP-checking. *Linguistic Inquiry* 16. 491–539.

Anand, Pranav. 2006. *De de se*. PhD dissertation. MIT, Cambridge, MA.

Andersen, Torben. 1999. Anti-logophoricity and indirect mode in Mabaan. *Studies in Language* 23.499–530.

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Baker, Mark. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 351–374. Amsterdam: John Benjamins.

Baker, Mark. 2018. Towards a shared syntax for shifted indexicals and logophoric pronouns. In Laura Kalin, Ileana Paul & Jozina van der Klok (eds.), *Heading in the Right Direction: Linguistic Treats for Lisa Travis*, 11–22. Montreal: *McGill Working Papers in Linguistics*.

Barbosa, Pilar P. 2009. Two kinds of subject *pro*. *Studia Linguistica* 63. 2–58.

Barbosa, Pilar P. 2019. *Pro* as a minimal nP: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50(3). 487–526.

Bianchi, Valentina. 2006. On the syntax of personal arguments. *Lingua* 116. 2023–2067.

Belletti, Adriana, Elisa Bennati, & Antonella Sorace. 2007. Theoretical and developmental issues in the syntax of subjects. *Natural Language and Linguistic Theory* 25. 657–689.

Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance languages*. Dordrecht: Foris.

Brattico, Pauli. 2017. Control and null subjects are governed by morphosyntax in Finnish. *Finno-Ugric Languages and Linguistics* 6. 2–37.

Bresnan, Joan. 1972. *Theory of Complementation in English Syntax*. Cambridge, MA: MIT PhD dissertation.

Cardinaletti, Anna, & Michael Starke. 1999. The typology of structural deficiency. In Henk van Riemsdijk (ed.), *Clitics and Other Functional Categories in European languages*, 145–233. Berlin: Mouton de Gruyter.

Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. PhD dissertation. University of Massachussetts, Amherst, MA.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in language*, 1–52. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Clements, George N. 1975. Logophoric pronouns in Ewe. *Journal of West African Languages* 10. 141–177.

Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32. 1055–1094.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33. 409–442.

Demirdache, Hamida, & Myriam Uribe Etxebarria. 2005. Aspect and temporal modification. In Paula Kempchinsky & Roumyana Slabakova (eds.), *Aspectual Inquiries*, 191–221. Dordrecht: Springer.

Diercks, Michael. 2013. Indirect Agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31. 357–407.

Duarte, M. Eugênia L. 1995. *A perda do princípio "evite pronome" no Português Brasileiro* [The Loss of the Avoid Pronoun Principle in Brazilian Portuguese]. Campinas, São Paulo: State University of Campinas PhD dissertation.

Embick, David, & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32. 555–595.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistic Theory* 25. 691–734.

Giorgi, Alessandra. 2009. A Grammar of Italian Sequence of Tense. *University of Venice Working Papers in Linguistics* 19. 111–156.

Giorgi, Alessandra, & Fabio Pianesi. 2004. Complementizer deletion in Italian. In Luigi Rizzi (ed.), *The Structure of CP and IP. The Cartography of Syntax, Volume 2,* 190–210. New York: Oxford University Press.

Grano, Thomas. 2012. *Control and restructuring at the syntax-semantics interface*. PhD dissertation. University of Chicago, Chicago, IL.

Grimshaw, Jane & Vieri Samek-Lodovici. 1996. Optimal subjects. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, 589–605. Amherst, Mass: GLSA.

Halle, Morris, & Alec Marantz. 1993. Distributed Morphology. In Kenneth Hale & Samuel J. Keyser (eds.), *The View from Building 20*. Cambridge, MA: MIT Press.

Higginbotham, James. 2001. Why is sequence of tense obligatory? *Oxford University Working Papers in Linguistics, Philology & Phonetics* 6. 67–90.

Holmberg, Anders. 2005. Is there little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010a. The null generic subject pronoun in Finnish: A case of incorporation in T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders. 2010b. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders, & Michelle Sheehan. 2010. Control into finite clauses in partial null subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–152. Cambridge: Cambridge University Press.

Huang, C.-T. James. 1989. *Pro*-drop in Chinese: A generalized control theory. In Osvaldo Jaeggli & Ken Safir (eds.), *The Null Subject Parameter*, 182–214. Dordrecht: Kluwer.

Jaeggli, Osvaldo, & Kenneth Safir. 1989. The null subject parameter and parametric theory. In Osvaldo Jaeggli & Ken Safir (eds.), *The Null Subject Parameter*, 1–44. Dordrecht: Kluwer.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2. 39–76.

Jelinek, Eloise. 1998. Voice and transitivity as functional projections in Yaqui. In Miriam Butt and Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, 177–206. Stanford: CSLI

Kempchinsky, Paula. 2009. What can the subjunctive disjoint reference effect tell us about the subjunctive? *Lingua* 119. 1788–1810.

Kiparsky, Paul, & J.-Frits Staal. 1969. Syntactic and semantic relations in Panini. *Foundations of Language* 5. 83–117.

Kiparsky, Paul. 2002. The architecture of Panini's grammar. Three lectures delivered at the Hyderabad conference on the architecture of grammar and at the University of California, LA. Ms. http://www.stanford.edu/~Papershyderabad.pdf. Date of access: 2020/07/26.

Koopman, Hilda, & Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20. 555–588.

Krapova, Iliyana. 2001. Subjunctives in Bulgarian and Modern Greek. In María Luisa Rivero & Angela Ralli (eds.), *Comparative Syntax of Balkan Languages*, 105–126. Oxford: Oxford University Press.

Landau, Idan. 2000. *Elements of Control*. Dordrecht: Kluwer.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22. 811–877.

Landau, Idan. 2015. *A Two-tiered Theory of Control*. Cambridge, MA: MIT Press.

Maling, Joan. 1984. Non-clause bound reflexives in Modern Icelandic. *Linguistics and Philosophy* 7. 211–241.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2015. How much context is enough?: Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46. 273–303.

Modesto, Marcelo. 2007. Null subjects in Brazilian Portuguese and Finnish: They are not derived by movement. In William D. Davies, Gregory Stump & Stanley Dubinsky (eds.), *New Horizons in the Analysis of Control and Raising*, 231–248. New York: Springer.

Neeleman, Ad & Kriszta Szendrői. 2007. Radical *pro*-drop and the morphology of pronouns. *Linguistic Inquiry* 38. 671–714.

Ogihara, Toshiyuki. 2007. Tense and aspect in truth-conditional semantics. *Lingua* 117: 392–418.

Partapuoli, Irene. 2020. Davvisámegiela Persovnnalaš Pronomenat Goalmmát Persovnnas. [Third Person Pronouns in North Saami.] Umeå university term paper.

Patel-Grosz, Pritty, & Patrick G. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48. 259–297.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linquistic Inquiry* 20. 365–424.

Reichenbach, Hans. 1947. The tenses of verbs. In Maria Reichenbach (ed.), *Elements of symbolic logic,* 287–298. New York: The Macmillan Company.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–558.

Roberts, Ian. 2010. A deletion analysis of null subjects. In Theresa Bieberauer, Anders Holmberg, Michelle Sheehan & Ian Roberts, *Parametric Variation: Null Subjects in Minimalist Theory*, 58–87. Cambridge: Cambridge University press.

Rohrbacher, Bernard. 1999. *Morphology-driven syntax: A Theory of V to I Raising and Pro-drop*. Amsterdam: John Benjamins.

Safir, Ken. 1985. Subject clitics and the nom-drop parameter. In Hagit Borer (ed.), *Clitics in Syntactic Theory*, 333–356. New York: Academic Press.

Safir, Ken. 2004. Person, context and perspective. *Rivista di Linguistica* 16. 107–153.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality-theoretic Analysis*. PhD dissertation. Rutgers University, New Brunswick, NJ.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26. 29–120.

Scheller, Elisabeth & Mikael Vinka. 2016. The Saami languages. In Olsson Mats-Olov (ed.), *The Barents Encyclopedia*, 265–270, Oslo: Pax Forlag.

Shushurin, Philip. 2017. Null pronouns in Russian embedded clauses. In Pritty Patel-Grosz, Patrick Georg Grosz & Sarah Zobel (eds.), *Pronouns in Embedded Contexts at the Syntax/ Semantics Interface*, 145–169. Berlin: Springer.

Sigurðsson, Halldór Ármann. 1993. Argument-drop in Old Icelandic. *Lingua* 89. 247–280.

Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense, and speech features. *Revista di linguistica* 16. 219–251.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42. 267–304.

Sigurðsson, Halldór Ármann. 2014a. About pronouns. *Working Papers in Scandinavian Syntax* 92. 66–98.

Sigurðsson, Halldór Ármann. 2014b. Context-linked grammar. *Language Sciences* 46. 175–188.

Sigurðsson, Halldór Ármann, & Joan Maling. 2008. Argument drop and the empty left edge condition. *Working Papers in Scandinavian Syntax* 81. 1–27.

Spyropoulos, Vassilios. 2007. Finiteness and control in Greek. In William D. Davies & Stanley Dubinsky (eds.), *New Horizons in the Analysis of Control and Raising*, 159–183. Dordrecht: Springer.

Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13. 561–570.

Terzi, Arhonto. 1992. Pro *in finite clauses. A study of the inflectional heads of the Balkan languages*. PhD dissertation. New York University, New York.

Terzi, Arhonto. 1997. Pro and null case in finite clauses. *The Linguistic Review* 14. 335–360.

Vainikka, Anne, & Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17. 613–671.

Vinka, Mikael. 2007. Verb movement and the morphosyntax of negation in Saami. In Erin Bainbridge & Brian Agbayani (eds.), *Proceedings of the Thirty-Fourth Western Conference on Linguistics*, 456–465. Fresno: Department of Linguistics California State University.

Vinka, Mikael. 2018. Pronouns and logophoricity in South Saami. In Trevor Driscoll (ed.), *Proceedings of the Thirtieth Western Conference on Linguistics*, 184–195. Fresno: Department of Linguistics California State University.

Vinka, Mikael. 2019. Pronouns, logophoricity and point of view in South Saami. Poster presented at the Chicago Linguistic Society 55th annual meeting.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11: 203–238.

Wurmbrand, Susi. 2014. Tense and aspect in English infinitives. *Linguistic Inquiry* 45. 403–447.

Zagona, Karen. 2014. Sequence-of-tense and the features of finite tenses. *Nordlyd* 41. 261–272.

Susann Fischer and Beáta Wagner-Nagy

12 Null subjects in Selkup and Nganasan

1 Introduction

The aim of this chapter is to provide an inventory of null subjects in Selkup and Nganasan.[1] On the basis of newly assembled corpus data (Brykina et al. 2018a, 2018b; Budzisch et al. 2019; see also Section 2.1 herein), we will discuss the conditions under which null subjects are licensed. We will also focus on the arbitrary readings of 3PL and 3SG (null and overt) pronominals in Selkup and Nganasan, with the ultimate goal of providing evidence that they are neither clearly consistent Null Subject Languages nor clearly partial Null Subject Languages.

As Holmberg explains in the Preface to this volume, there are four major types of Null Subject Languages, henceforth NSLs (see also Roberts & Holmberg 2010 for a detailed discussion). Thus, consistent NSLs include, for instance, Spanish and Greek. Languages like German or Dutch, on the other hand, belong to expletive NSLs. In turn, radical *pro*-drop or radical NSLs embrace Chinese and Indonesian, among others. Finally, languages like Finnish and Russian are classified as partial NSLs. For each type, different criteria have been established in order to classify the existing null subject languages of the world.

In the meantime, a lot of research has been devoted to proving that Roberts & Holmberg's typological classes should not be seen as discrete types but rather as a continuum. For instance, Hungarian, which belongs to the Finno-Ugric language family, has been convincingly argued by Dalmi (2014a, b) to be a NSL that does not fully fit in any major type of Roberts & Holmberg's typological system. Thus, in order to understand what type of NSLs Nganasan and Selkup represent, in section 2 we will briefly compare consistent NSLs, also known as classical NSLs (e.g. Rizzi 1986; Manzini & Savoia 1997; Borer 2001: 176) with partial NSLs. There is no need to compare Nganasan and Selkup to expletive NSLs, which only sometimes drop expletives, or to radical NSLs, which are of the isolating type concerning their morphology.

A quick glance at the two languages in question allows us to observe that they are of the agglutinative morphological type and, as will be shown in sections 3 and 4, they drop referential and expletive subjects in all types of sentences. Thus, sections 3 and 4 provide an overview of Selkup and Nganasan, respectively. Section 5 concludes our chapter.

---

2 A brief overview of Selkup and Nganasan

The languages under discussion here belong to the Samoyedic branch of Uralic languages. Selkup, which is in older sources referred to as Ostyak-Samoyed, is spoken in Central-Western Siberia along the tributary rivers of Yenisei and Ob in the Tomsk Region, the Yamal-Nenets Autonomous Area, and in the Turukhansk district of the Krasnoyarsk Territory. Selkup is fractured into a large number of dialects or even a dialect continuum which can be divided into at least three dialectal, largely diversified groups. The dialectal groups and dialects are as follows: the Northern group (Middle and Upper Taz, Yelogui, Upper Tolka and Baikha dailects), the Central group (Tym, Narym, Vasyugan dialects) and the Southern group (Middle and Upper Ob, Ket, Chulym and Chaya dialects).[2] In the 2010 Russian census, 3,527 individuals identified themselves as ethnic Selkups. Although Selkup has slightly more speakers (approx. 1000) than Nganasan, it is also on the verge of extinction. The majority of the speakers or semi-speakers use one of the varieties of the Northern dialectal group. Neither the Central dialects nor the Southern ones count more than 15–20 active speakers. The Chaya and Chulym dialects have already disappeared. All speakers of Selkup are bilingual, also speaking Russian. In the North they are even trilingual, i.e. they speak Selkup, Russian and Evenki or Ket.

Nganasan (in older sources referred to as Tavgi-Samoyed) is spoken in the Northern part of Siberia, on the Taimyr Peninsula. Presently, there are only approx. 100 people who speak Nganasan. According to the latest Russian census (2010), the number of ethnic Nganasan is 862, but in reality, there are far fewer people who still speak the language. Almost all Nganasans speak Russian. The use of Nganasan is restricted to a few villages in the Taimyr Dolgan-Nenets Municipality District. Most of the Nganasan people live in Usť-Avam and in Volochanka.

Both Selkup and Nganasan are nominative-accusative alignment languages, showing the usual distribution of argument marking in transitive and intransitive sentences. Word order is predominantly SOV, with possible exceptions attributable to the information structure and other pragmatic issues. Both languages belong to the agglutinative synthetic language type with suffixation as the dominant morphological process. However, there are quite a few differences between Nganasan and Selkup. In Nganasan we also find inflectional properties, which can be seen from the fact that number and case suffixes are not as clearly

---

2 For more information about the dialects, see Janurik (1978) and Glushkov (2007, 2013), among others.

separable as in Selkup. Selkup has prefixes, Nganasan does not. Selkup has thirteen different cases, Nganasan seven.

The noun has three numbers (singular, dual, and plural) in both langues. The verb agrees in person and number with the subject and the direct object. It also shows two types of conjugation in Selkup, i.e. subjective and objective, while in Nganasan there is the third type, i.e. the reflexive conjugation. All these features interact in the use of null subjects in these languages.

2.1 About the data

While the study presented here is predominantly corpus-based, older grammatical descriptions are also taken into account. The Selkup examples are taken from a large number of spoken sources containing mostly folklore text, but a smaller portion of these texts are everyday narratives. Examples coded as [XYZ_ year_title of text] refer to texts from the Selkup Language Corpus[3] (SLC) (marked by d_, see Budzisch et al. 2019), which consists of already published texts from the Northern, Central, and Southern dialects. The published version of SLC contains 144 transcripts with 9,156 sentences (55,839 tokens) from 53 speakers. This includes 26 transcripts in Northern Selkup, 48 in Central Selkup, 66 in Southern Selkup, and 4 in a mixed dialect (Central/Southern Selkup). The other source is the INEL Selkup Corpus[4] (marked by a_, see Brykina et al. 2018a). This corpus is composed of texts from the archive of Angelina Ivanovna, who collected valuable materials between 1962 and 1977. She gathered a large amount of the Selkup data in almost all regions where the Selkup people lived. The published corpus contains 78 transcripts (30 folklore, 40 narrative, 7 translations, 1 song) of 47 speakers with 3,273 sentences (18,673 tokens). Most of the transcripts (57 texts) originate from the speakers of the Northern dialect group, 7 from the speakers of the Central dialect group (Tym, Narym), and 14 texts from the speakers of Ket dialect.

The Nganasan data are taken from the Nganasan Spoken Language Corpus (Brykina et al. 2018b).[5] The materials of this corpus come from more than 13 researchers as well as from the colleagues of the Taimyr Radio. The transcripts of

---

3 The compilation of this corpus has been supported by the German Research Grant.
4 The corpus has come into being in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Program, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Program is coordinated by the Union of the German Academies of Sciences and Humanities.
5 The compiling of the corpus has been supported by the German Research Grant.

the corpus come from 57 speakers and consist of 176 transcripts with 22,126 sentences (143,993 tokens), as well as approx. 32-hour long audio recordings.

2.2 The finite verb and its arguments in Selkup and Nganasan

In both languages, the finite verb is structurally built as follows: stem + (derivational suffixes) + tense marker or co-affix + mood marker + personal ending. In Selkup, depending on the phonological structure of the verb, the personal endings and some mood markers such as imperative and conditional are connected to the stem with a co-affix (linking element), e.g. Middle Ob *tü:-wa-n* [come-co-3sg] but Middle Taz *ila-k* [live-1sg] 'I live.' In Nganasan, however, the use of the co-affix is obligatory. Selkup distinguishes two conjugation systems: subjective and objective, whereas in Nganasan, a reflexive conjugation also occurs. Both the subjective and the objective conjugation types have their own set of inflectional endings. Personal suffixes of the subjective conjugation express the person of the subject, while the suffixes of the objective paradigm refer not only to the subject but indicate the object as well, e.g. Tym Selkup *čo:nda-p* [dress-1sg>sg]'I dress him/her', Chaya Selkup *pajaga ťepbɨ-ni-n* [old.woman dress-co-3sg] 'The old woman got dressed.'

3 Licensing lexical and null subjects

3.1 Features of null subjects in consistent and partial NSLs: A short overview

Consistent NSLs are languages that allow referential pronominal subjects (1a), expletives (1b) and quasi-arguments (1c)[6] to be null or unexpressed in any person-number combination, in any tense, in any type of sentence, i.e. in *wh*-questions, matrix clauses and embedded clauses.[7]

---

6 We will not make any difference between null quasi-arguments and expletives, but call both *expletives*. Even though in expletive NSLs – like, e.g., German – only true expletives can be dropped, whereas quasi-arguments can never be dropped.

7 Unless otherwise indicated, all the Spanish examples were provided and checked by native speakers.

(1)   a.   *Hablo*   *español.*   Referential null subject
              speak.PRS.1SG   Spanish
              'I speak Spanish.'

       b.   *Hay*   *un perro en el*   *jardín.*   Expletive null subject
              has.PRS.3.SG   a   dog   in   the   garden
              'There is a dog in the garden.'

       c.   *Llueve.*   Null quasi-argument
              rain.PRS.3.SG
              'It rains.'
              (Spanish)

In the traditional generative literature, several additional criteria are discussed that seem to correlate with being a consistent NSL. We will mainly focus on two of them here.[8] It has been pointed out that in NSLs overt subjects in embedded contexts, in contrast to non-NSLs like English or French, do not normally allow the interpretation in which the overt pronoun refers to the subject in the main clause (Roberts & Holmberg 2010: 7), consider (2). The pronoun *él* of the embedded clause strongly prefers the interpretation which is disjoint from *el professor* 'the professor' in the main clause.

(2)   *El*   *professor habló después que (él)*   *llegó.*
       the   professor   spoke   after   that   (he)   arrived
       'The professor spoke after he arrived.'
       (Spanish)

Furthermore, it has been shown that generic pronouns in a consistent NSL seem to require a special overt element (clitic) or verb-form, (3). In languages allowing any referential subject to be dropped freely, the default reading of a dropped 3SG subject is referential 3SG (he/she/it) and not generic (see Holmberg & Sheehan 2010). Thus, in order to trigger the generic reading, an overt element, e.g., *se* in Spanish, needs to be included in the sentence.

---

8 There are other syntactic criteria argued to correlate with the NSL status, e.g. subject inversion, the definiteness effect, and the *that-t effect*. We will not discuss these criteria here, but refer the reader to Alexiadou & Anagnastopoulou (2007), Fischer (2012, 2016), Perlmutter (1971) and Rizzi (1986), among many others. The criteria mentioned are without any doubt important, however, they seem more important for generative theorizing and the status of parameters than for deciding to which type of NSL a given language belongs.

(3) *Aquí no se puede fumar.*
here not SE can smoke
'One can't smoke here.'
(Spanish)

As for partial NSLs, they allow some referential null subjects but not others. Finnish, for example, allows 1st and 2nd person to be null, whereas 3rd person referential pronouns can only be omitted when bound by a higher argument, under conditions that are rather poorly understood (cf. Roberts & Holmberg 2010: 11), compare examples in (4) with example (5).

(4) a. (*Minä*) *puhu-n englanti-a.*
(I) speak-1SG English-PART
'I speak English.'

b. (*Sinä*) *puhu-t englanti-a.*
(you) speak-2SG English-PART
'You speak English.'

c. *(*Hän*) *puhu-u englanti-a.*
(s/he) speak-3SG English-PART
'S/he speaks English.'

d. (*Me*) *puhu-mme englanti-a.*
(we) speak-1PL English-PART
'We speak English.'

e. (*Te*) *puhu-tte englanti-a.*
(you.PL) speak-2PL English-PART
'You speak English.'

f. *(*He*) *puhu-vat englanti-a.*
they speak-3PL English-PART
'They speak English.'
(Finnish, Holmberg 2005)

(5) *Pekka$_i$ väittä-ä* [*että hän$_{i/j}$/pro$_{i/*j}$ puhu-u englanti-a hyvin*].
Pekka claim-3SG that s/he speak-3SG English-PART well
'Pekka claims that he speaks English well.'
(Finnish, Holmberg 2005: 539)[9]

---

9 The glossing and the free translation are our own and slightly different from Holmberg's (2005: 539).

What we said with respect to the default readings in consistent NSLs can be applied here as well. In Finnish, 3SG covert subjects (*he, she, it*) need an antecedent in order to be interpreted as referential, as in (5). Consequently, any 3SG covert subject that has no antecedent must be interpreted as generic by default, as in (6):

(6) *Tä-ssä   tuoli-ssa   istu-u   mukavasti.*
     this-INE   chair-INE   sit-3SG   comfortably
     'One[10] sits comfortably in this chair.'
     (Finnish, Holmberg 2010: 93)

In order to see what type Nganasan and Selkup belong to, or where exactly they can be placed in the continuum of the proposed types, we will discuss what kind of sentences allow null subjects and take a brief look at generic overt and covert pronouns and the readings they trigger in the two languages.

3.2 Licensing null subjects in Selkup

The subject of a transitive and intransitive sentence can be overt or covert in all types of sentences and in all tenses. The data from the SLC (Budzisch et al. 2018) show that in 46% of the cases, subjects are not expressed lexically in the sentence, the relevant numbers indicated in Figure 1 below.

The lexical subject appears in the nominative case and tends to take the sentence-initial position. The vast majority of overt subjects are expressed by a nominal phrase, which can be a common noun or a proper noun,(8a-c). Numerals and demonstratives are rarely used as subjects, however, some examples can be found (7). These are included in the nominal subjects in statistics (see Figure 1). As Figure 1 shows, only a small number of subjects are expressed by pronouns, (8d-e).

(7) *i   **šede-wi**   qʷaja-nʒa-j   tʷä:r-lä.*
     and   two-POSS.1PL   go-FUT-1DU   steal-CVB
     'The two of us will go to steal.'
     (Middle Ob dialect, [a/d_TFF_1967_FedjkaVillageThief_flk.176])

---

10 In Finnish the difference between generic inclusive vs. generic exclusive is expressed by different structures. Sentence (6) includes the speaker and addressee in its reference, for an exclusive reading a passive construction is used (cf. Holmberg 2010: 203). See Section 3 of this chapter.
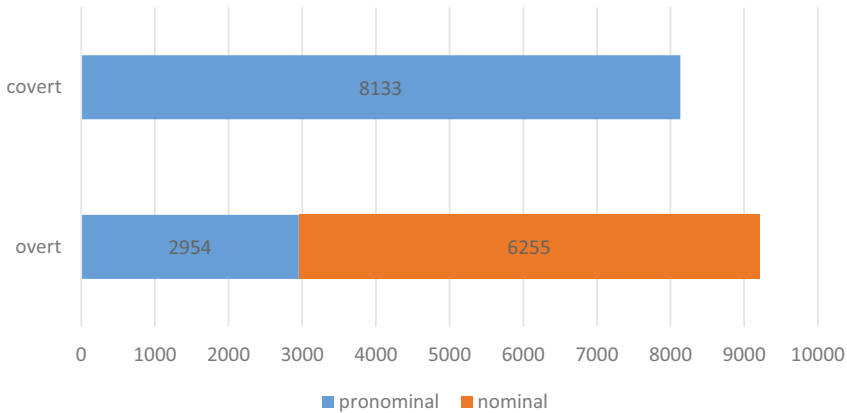
Figure 1: The proportion of covert and overt subjects in Selkup.

(8) a. *aŋgə   wašešpi-nda.*
       capercaillie   fly-INFER.3SG
       'The capercaillie seems to be flying.'
       (Tym dialect, [a/d_JIF_1968_Kamacha_nar.009])

    b. *Pönəge   qʷäl   tʷer-na.*
       Pönege   fish   steal-CO.3SG
       'Pönege steals the fish(es).'
       (Tym dialect, [a/d_KFN_1967_BigBear_flk.002])

    c. ***päläľ   topɔː-ľ   qəńarɔːľ   tɪmńa-tɔː***   *päläľ   utɔː-sä*
       half   leg-ADJ   crooked   brother-POSS.3SG   half   hand-INS
       sotta,   päläľ   utɔː-sä   čoqotɔː-ŋa.
       scoop.3SG   half   hand-INS   poke-CO.3SG
       'His one-legged crooked brother scoops with one hand, with another
       he hits.'
       (Taz dialect, [d_MG_1941_IchakichaAndGrandmother_flk.027])

    d. ***mat***   *tap-čeːl   qʷaja-ɣa-k   poq-ɨ-la-p   mannɨmbɨ-gu.*
       I   this-day   go-CO-1SG   net-EP-PL-ACC   look-INF
       'I went to see the net today.'
       (Narym dialect, [a_IAI_1968_Fishing_nar.001])

e. *nɨnɨ* [*loːsɨ ira*]ᵢ   *mɔːt-tɨ šeːr-ľä   puːľa täp*ⱼ
    then devil old.man tent-ILL come.in-CVB after s/he
    pona   tanta
    outwards go.out.3SG
    'Then, after the old devil went into the house, he went out into the street.'
    (Taz dialect, [a_KAI_1965_BoyAndOldDevil1_flk.228])

Sentence (8e) shows a pronoun in the embedded clause which is not co-indexed with the full DP in the matrix sentences (it is disjoint from 'the old devil'). This is exactly what we expect from consistent NSLs, in contrast to non-NSLs (cf. Roberts & Holmberg 2010).

All persons can be covert in all sentences across all tenses. Both referential and expletive subjects can be freely dropped,[11] as is the case in consistent NSLs, according to the typology of Roberts & Holmberg (2010). The subject is identified by personal verbal suffixes, in the sense of Rizzi (1982). First person and second person pronominal subjects are freely dropped in any context. The 3ʳᵈ person subject is usually dropped when it has been expressed earlier in the discourse, and also when the speaker assumes that the addressee already knows who s/he is talking about. Thus, unlike in Finnish, which is a partial NSL, it does not need to be bound by a higher argument (i.e. an antecedent in the main clause), it can be bound in the discourse. The examples in (9a)–(9c) serve to illustrate a first, a second, and a third person pronoun, which has been dropped, respectively. In turn, (9d) shows a dropped 3SG referential subject in the embedded clause, which differs in reference from the dropped 3SG in the main clause. Finally, (9e) presents a dropped subject in a *wh*-question.

(9) a. pro *nɨmtɨ učittɨ-sa-k   sompola poː-t   kuntɨ.*
         (I)   there learn-PST-1SG five   year-GEN during
         'I studied there for five years.'
         (Middle Taz dialect, [a_SMV_196X_Lifestory_nar.007])
     b. pro   *innä ši i-ŋa-ntɨ.*
         (you) up   me take-CO-2SG
         'You've pulled me up.'
         (Upper Taz dialect, [a_KAI_1965_OldManWithLittleMind1_flk.025])

---

11 We use the term 'dropped' and will include *pro* in the examples used. *Pro* will be located in the position where the overt subject would be located in an SOV order. However, at this point we cannot yet decide whether for Nganasan and Selkup a projected subject position for all sentences needs to be postulated.

c. pro   *maziŋ*   *so:rriku-s.*
(he) me.ACC love-PST.3SG
'He loved me.'
(Middle Ket dialect, [a_KMS_1963_Lifestory1_nar.019])

d. pro₁ *käti-sä-t*   pro₂ *soma* *ɛ:-ja*
(he) say-PST-3SG>SG   (it)   good   be-CO.3SG
'He said [that] it is good.'
(Middle Taz dialect, [a_AR_1965_RestlessNight_transl.029])

e. pro   *qaj-sä*[12]   *šäqqɛ-ntɔ:-min?*
(we) what-INS   spend.night-FUT-1PL
'What do we spend the night with?'
(Middle Taz dialect, [d_APA_1977_Kitten_flk.075]

As mentioned before, in consistent NSLs all pronouns regardless of the person and tense can be null (cf. Holmberg 2010). Examples (9a) and (9c) feature a null subject occurring with a verb in the past tense, (9b) in the present tense, and (9e) in the future tense, respectively. Selkup allows referential null subjects (10), expletive null subjects, (11), and generic exclusive null subjects, (12). These will be scrutinized shortly.

(10) pro_i *čap*   *tü-sa*   *mɔ:t-qin-ti,*   pro_i *qo-si-ti*
(she) hardly come-PST.3SG house-ILL-POSS.3SG (she) see-PST.3SG>SG
pro_j *illə*   *ɔ:l̓čimpa*
(it)   down fall.down.3SG
'As she came to her house, she saw that it had been destroyed.'
(Middle Taz dialect, [a_SAlAn_1965_Soldatka_nar.012])

(11) pro_EXPL *lipqimɔ:ta-ja-**n***
(it)   get_dark-CO-3SG
'It got dark.'
(Middle Taz dialect, [a_ IF_196X_WomanAndDevil_flk.005])

Expletive null subjects are typically used with *weather*-verbs or verbs describing natural phenomena (12a-d). These verbs can be basic (i.e. non-derived) or derived. An example of a basic verb is the Middle Taz verb *sə:riqo* 'to rain', while an example for a derived verb is Middle Taz *taŋimqo* 'to become summer'. Most of the derived verbs are formed with the translative suffix *(-m)* or with the subject

---

12 *Wh*-pronouns in Selkup and Nganasan are always in situ.

-*mɔ:t*. The latter suffix is normally attached to bound stems as in Middle Taz *lɨpkɨ-mɔ:t-qo* 'to get dark' (see (11) above and (12d) below). These verbs never take overt arguments, i.e. neither a lexical nor a suppletive pronominal subject can occur in the sentence. The only verbal ending attested is third person singular, and they never appear in the causative or imperative form.

(12) a. pro   *ɨllä*   *taŋɨm-nɨ.*
      (EXPL)   down   get_summer-CO.3SG
      'It is getting summer.'
      (Middle Taz dialect, [d_AVA_1973_Ichakicha_flk.095])

     b. pro   *aqqol*   *serro-ńi.*
      (EXPL)   again   rain-CO.3SG
      'It rains again.'
      (Upper Ket dialect, [d_KKN_1971_IvanAndGod_flk.020])

     c. pro   *ta:r*   *assə*   *ťe:lɨmbɨ-s,*   *ä:wɨ-t*   *qʷan-nɨ*
      EXPL   still   NEG   get_day- PST.3SG   mother-POSS.3SG   go_away-CO.3SG
      nä:-m-də   *per-gu.*
      daughter-ACC-POSS.3SG   look_for-INF
      'It didn't even get light (it wasn't even day), the mother went to look
      for the daughter.' [a_KMS_1963_BearAteTwoWomen1_nar.021]

     d. pro   *uže*   *lɨpqiɔːtpa,*   *me*   *oš*   *onɨt*   *pɛlɨkɔːlɛ:-ŋɔ:-mɨn.*
      EXPL   already   get_dark.3SG   we   EMPH   ourselves   stay_alone-CO-1PL
      'It is getting dark, obviously we were left alone.'
      (Taz dialect, [a_IF_196X_WomanAndDevil_flk.009])

Complement clauses are regularly attested in object position and they are predominantly finite, but (as we have seen above), there are complement clauses with non-finite forms, as well (12c). In clauses in which the subject is different from the subject of the matrix clause, the subject must be expressed by a lexical NP and the verb of the complement clause is inflected for subject and object agreement, as in (13a). If the subject is the same, either the so-called *supine form* or the infinitive form is used.[13] The choice of the two forms might be ideolectal. However, it should be noted that the supine form is only used in the Northern dialect. In sentences with the supine form, the subject of the complement clause is obligatorily coded by the possessive suffix. Additionally, the corresponding pronoun can appear in the genitive case preceding the nominalized verb, as shown in (13b). In sentences with the infinitive form, the subject cannot be marked.

---

13 The supine (or purposive converb) is used to form purpose clauses.

(13) a. ... *qajno mekga ära-u* *ťäri-n što struška*
    why me.DAT husband-POSS.1SG say-3SG that scobs
    üt-tə papalan ik azə
    water-ILL how IMP.NEG become.3SG
    'Why did my husband say that the scobs shouldn't fall into water?'
    (Chaya dialect, [a_PVD_1964_UnfaithfulWifeAndRobbers_flk.031])

 b. *äsa-ľ tinta nıľčik top-sı-tı* [*šitimtaliľ*
    father-POSS.2SG you.DAT so speak-PST-3SG>SG second
    *pit **tan** šäqı-**qunta:qa**].
    at.night you.GEN spend.night-SUP.POSS.2SG
    'The father told you to spend the second night.'
    (Middle Taz dialect, [a_NEP_1965_FoolInSackCoat_flk.056])

 c. ... *ta ľenťaj-ŋɔ:-nt, čajnik asa kika-ntı* [*musirı-**qo***].
    you lazy.person-CO-2SG teapot NEG want-2SG cook-INF
    'You are lazy, you do not want to boil the tea water.'
    (Baikha dialect, [a_PM_1977_IchakechikaAndQolsaqo_flk.021])

In impersonal sentences with an arbitrary reading, the third person plural form of the verb must be used. There is no overt lexical subject in (14).

(14) *ä:zı-n mü:ttä-ɣın täb-ı-m qʷät-pa-ttə.*
    father-GEN war-LOC s/he-EP-ACC kill-PRF-3PL
    '(They) killed him in World War II (lit. fathers war).'
    (Middle Ket dialect, [a_KMS_1963_Lifestory1_nar.021])

One of the most interesting topics related to null subjects in these languages is how dropped 3SG/3PL subjects are interpreted as generic. It has been argued that in consistent NSLs overt elements must be used to obtain the generic reading. In addition, it has been discussed controversially whether these overt elements (e.g. *on* in French, *man* in German, *se* in Spanish, *si* in Italian) should cover the inclusive vs. exclusive generic readings (see D'Alessandro & Alexiadou 2002; Holmberg 2005), the generic, the arbitrary, or the specific readings (see Egerland 2003), the quasi-universal or quasi-existential readings (see Cinque 1988), or, alternatively, whether the respective readings require different markers in a language.[14]

Selkup generic null subjects can have either the inclusive or the exclusive interpretation. 3PL null subjects are normally interpreted as generic exclusive (see

---

14 For the purposes of this chapter, we will use the terms inclusive and exclusive to refer to the readings including or excluding the speaker.

(15)). The speaker uttering the sentence in (15) is not from Ivankino. For the inclusive reading of 3PL, a lexical antecedent in the left-adjacent clause is needed. The passage given in (16a)–(16c) shows that the generic interpretation of the embedded 3PL null subject in (16b) is dependent on the overt DP *qu-la* (person). 3SG in (16c) has a generic reading referring to people in general including the speaker and the addressee.

(15) *Iwankina-ɣən arɨ-n   kuluba-tte*
      Ivankino-LOC   other-ADV speak-3PL
      'People speak otherwise in Ivankino.'
      (Chaya dialect, [a_PVD_1964_DialectInIvankino_nar.001])

(16) a. *aʒä-m   čaʒe-ndel   qup   e:-ɣa.*
         father-POSS.1SG drive-PTCP.PRS person   be-CO.3SG
         'My father is a chauffeur.'
         (Narym dialect, [KoIA_2019][15])
      b. *čenči-ɣa   niľʒik: **qu-la**   mi-nan   tänhe   uʒä-t,*
         speak-CO.3SG so   person-PL   we-ADE   reliable   work-3PL
         nemdi   pjaldïmba-dït, čaʒe-fa-tko   fa-ŋ   mannïmba-t,
         each.other   help-3PL   drive-ACT-TRL   good-ADV   see-3PL
         arkak   čaʒenda-t.
         carefully   drive-3PL
         'He said: our people are reliable, we (lit. they) always help each other, and we (lit. they) follow the cars and drive carefully.' [KoIA_2019]
      c. *kuʒat kod-ta   üte-ʒi-t –   čaʒ-le   a: omdešpa*
         when   who-DEF drink-FUT-3SG>SG   drive-CVB   NEG   sit.3SG
         'If one drinks, one does not drive.'
         [KoIA_2019]

The generic exclusive reading can be identified by the 3PL verbal ending in (14) and (15), and by the generic DP 'people' in (16b) and (17b). More specifically, according to Cabredo-Hofherr's (2003: 83) classification of possible arbitrary readings of 3PL null pronominals, (17a) exemplifies a universal reading which is licensed by the locative *Moscow*.[16]

---

15 These sentences are not from the corpus, but elicited by Alexandre Arkhipov from a speaker [KoIA] in an interview taken in 2019.

16 In Egerland's (2003) classification, Egerland's (2003) classification is a generic reading, yet for Cinque (1988) it would be quasi-universal. (17b) illustrates an inclusive generic reading according to D'Alessandro & Alexiadou's (2003) and a specific reading according to Egerland (2003).

(17) a. *Moskwa-ɣet   šerešpa-det kʷeẑedet kaborg.*
     Moscow-LOC   wear-3PL   beautiful cloth
     'People in Moscow wear beautiful clothes.'
     (Narym dialect, [KoIA_2019])

   b. ***qu-la,***   *kuʒat ütežla-də, mešpa-det koštel mɨ-p*
     person-PL when drink-3PL do-3PL   bad   thing-ACC
     'People/they do bad things when they drink.'
     [KoIA_2019]

In sum, we find all types of referential subjects across all types of sentences. There are generic subject constructions attested, which are either identified by the 3SG/3PL suffixes on the verb or by the lexical subject *qula* 'people'. The generic inclusive reading can only be obtained by using the lexical subject *qula* and by the 3PL suffix when bound by a lexical antecedent in the left-adjacent clause. The generic exclusive or the universal reading, which does not include the speaker, is obtained by the mere [3PL] suffix when no lexical antecedent in the left-adjacent clause binds it; the generic inclusive reading is affected by a 3SG suffix.

3.3 Licensing null subjects in Nganasan

In Nganasan, as in Selkup, subjects of intransitive and transitive sentences can be overt or covert in all types of sentences irrespective of the tense. Subjects in Nganasan are expressed by a DP composed of a modified noun (18a), a common noun, (18b), a proper noun (18c) or a pronoun (18d), nominative case-marked. There is no morphological marking to distinguish between the agentive vs. non-agentive or human vs. non-human subject. As noted above, the predicate must agree in person and number with the subject. Grammatical subjects tend to appear in sentence-initial position, however, they can be preceded by adverbs and question words.

(18) a. *təti ŋĩantu-rə   ma-tə-tu   čii-Ɂə,   siti*
     this young.man-POSS.2SG tent-LAT-POSS.3SG come.in-CO.3SG s/he
     ńemɨ-ntɨ   ma-tə.
     mother-GEN.POSS.3SG   tent-LAT
     'This young man went home into his mother's tent.'
     [JSM_090809_Life_nar.119]

b. *təgətə ŋonəə* **ďeďi-čiŋ** *mejkə-ti-ʔ əligaku-j*
then one.more father-PL.POSS.3PL make-CO-3PL small-PL.ACC
kəndaʔku-j. . .
small.sled-PL.ACC
'And then, their fathers used to make little sleds (for
them). . .' [ChND_041213_Reminiscence_nar.141]

c. *Śimbia ŋəðü-śüə Kuruʔmaku.*
Simbia see-PST.3SG Kurumaku
'Simbia saw Kürümaku.'
[ChND 2006][17]

d. **mənə** *sovetskəi-ndə ŋojbəuʔkə-suðə-m tiŋ na-nduʔ*
I Soviet-LAT work-FUT-1SG you.PL.GEN with-POSS.2PL
'I will work with you together as a Soviet man.'
[MDN_97_War1_nar.004]

Numerals (19a), participles (19b) and demonstratives (19c) are also used as
subject. A large part of the participles that are used as subjects are lexicalized
nouns expressing the names of professions (e.g. *basutuə* 'hunter'). In some cases,
non-lexicalized present participles can also appear as subjects (19b).

(19) a. *ŋamīajďüm-ə ďalɨ ŋonəə sʲajbə bii-ʔīa.*
other-EP.GEN day one_more seven go_away-CO.3SG
'Next day another seven left.'
[ChND_080729_SevenNjote_flks.093]

b. *ouʔ, tahañā mað-ə kunsɨnɨ* **buə-tuə** *sojbu-munu-ču.*
oh now tent-EP.GEN inside speak-PTCP.PRS sound-AU-POSS.3SG
'Someone is heard speaking inside the tent.' [JDH_00_
TwoDolganBrothers_flk.456]

c. *təti-rə bii-ʔīai-ðə ma-tu ďa*
that-POSS.2SG go.away-CO.3SG.RFL tent-GEN.POSS.3SG ALL
turimiď-i-ʔə bii-ʔīa.
shrink.back-EP-PRF.3SG go.away-CO.3SG
'He left, went back home.'
[KNT_960809_WildAnimals_flkd.164]

---

17 This sentence is elicited from the speaker [ChND].

The analysis of the data from a sub-corpus[18] of the Nganasan Spoken Language corpus (Brykina et al. 2018b) reveals that around 64% of all subjects are pronominal subjects, which are regularly omitted, since they can be identified (in the sense of Rizzi 1982, 1986) by verbal suffixes that mark person and number. Thus, only 12% of the pronouns are overt, as shown in Figure 2.
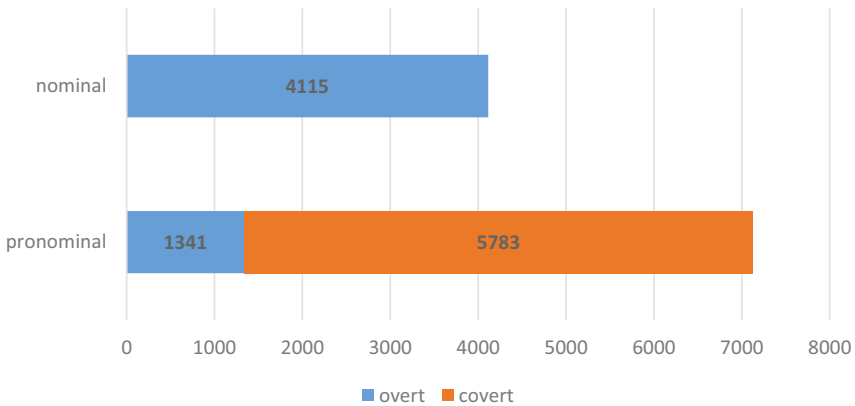
Figure 2: The relation between covert and overt subjects in Nganasan.

All subjects, i.e. referential pronominal, expletive and quasi arguments, across all persons and tenses, can be phonetically null. Examples (20a)–(20c) illustrate, respectively, the (dropped subject) 1SG, 2SG and 3SG pronoun. It should be noted that in some sentences the object is also covert and only the verbal ending refers to it, consider (20a). In turn, (20c) consists of two clauses; the first part of the sentence illustrates both the subject (*pro*_i) and the object (*pro*_j) as covert participants. In the second clause, the object of the verb of the first clause appears as the subject of the second clause. This participant, however, is still covert. The example in (20d) shows that the embedded *pro* (she) is not coindexed with the full DP (husband) in the main clause; *pro* is clearly bound in discourse.

(20) a. *təbtə* pro pro *ńi-siə-mə*   *ńaagəi-ʔ*   *dʼindi-ʔ*
         also   (1SG) (3SG) NEG.AUX-PST-1SG>SG good-ADV   hear-CNG
         'I did not hear it good.'
         [ChND-KES_061107_Berizenaa_flkd.027]

---

18 The sub-corpus consists of 11,239 sentences.

    b.   *maaðə*   pro   *əmə*   *ma-ðə-mtə*   *ńi-ĥīaði-**ŋ***

        why   (2SG) this.ACC   tent-DST-ACC.2SG   NEG.AUX-INFER-2SG

        melisi-ʔ

        make-CNG

        'Why haven't you put up this tent (for you)?'

        [ChNS_080214_Hibula_flkd.098]

    c.   pro$_i$ pro$_j$ *ńolhəda-ʔa-**ðu**,*   pro$_j$ *bii-ʔīai-**ðə**.*

        (3SG)   (3SG)   set.free-CO-3SG>SG   (3SG)   go.away-CO.REFL-3SG.REFL

        'He set it free, then it (the fox) went away.'

        [JSM_080212_Hibula_flkd.017]

    d.   *bəjkaʔa-tu*   *čii-ti-ndə,*   pro   *munu-ntu...*

        old.man-GEN.POSS.3SG   come.in-NMLZ-LAT   (3SG)   say-CO.3SG

        'After her husband entered the tent, she said. . .'

        [ChND_99_Barusi_flkd.033]

In addition, we also find expletive null subjects and quasi null subjects. In general, meteorological verbs do not have overt arguments. Only some temperature verbs, e.g. *čeśiśi* 'be cold', allow lexically expressed subjects. There are two types of *weather*-verbs, i.e. basic and derived. Most of the derived verbs are formed with the help of the translative derivational suffix *(-m)* from the noun such as *koðu* 'snowstorm' (21a), *hojməgəə* 'dark'. Furthermore, there are some weather nouns such as *ďüməðaʔa* or *ďüməðaŋku* 'snowfall' and *kourüðaŋku* 'sunny day' which usually take the predicate position and cannot have any arguments These nouns can be verbalized using the translative suffix (*-m*) as *ďüməðaʔamsa* 'become a snowy day' (see sentence (21b)).

(21) a.   pro   *kundu-śüði-ŋ*   *hiińďa*   pro

         (2SG) fall.asleep-FUT.REFL-2SG.REFL   at.night (3SG)

         ***kotum-sutə**.*

         get.snowstorm-FUT.3SG

         'When you will sleep, there will be a snowstorm.'

         [PKK_71_OneTent_flkd.059]

      b.   *ouʔ,*   *taharīā* *tə,*   *hojməgəə* *maa* *kuniðə*

         EXCL now   well dark   what from.where

         ŋəðə-tə-ŋu-mə   *hii-m-i-ʔə*   *tə,*

         find-FUT-INTER-1SG> SG night-TRL-EP-CO.3SG   well

         ŋonəə   ***ďüməðaʔi-m**-ü-ʔə-gəľiča.*

         one.more   get.snowfall-EP-CO-EMPH.3SG

         'Where would I find her, it's already dark and the snow falls.'

         [JDH_00_Njaakju_flkd.039]

There is a special sentence type expressing general truth in which the predicate is an adjective with a predicative suffix (the personal endings of the subjective conjugation). These are conditional sentences. The predicate adjective of the matrix sentence cannot have a subject, as it is always dropped. The subject of the subordinate sentence is the present form of the temporal-conditional converb, (22).[19]

(22) *nəŋhuə,   məu-ntuʔ   ləŋu-biʔ.*
      bad.3SG   earth-GEN.POSS.2PL burn-COND
      'The burning of your land is bad.'
      [KNT 1994][20]

Furthermore, we also find generic subjects with the exclusive and inclusive readings which can be covert or overt in Nganasan. The sentence in (23a) illustrates a third person exclusive null subject identified by the 3PL reading. Examples (23b) and (23c) show a lexical DP 'man/person', which, without any doubt, has a generic reading. We could interpret it as inclusive since in the context of the story it is clear that the woman talking in (23b) owns reindeers herself and the man in (23c) is talking about a group to which he belongs. The sentence in (23d) exemplifies a dropped 3SG which needs be interpreted as a generic construction according to Egerland (2003), or a quasi-universal according to Cinque (1988).

(23) a. *aniʔə   śitəbi-təni   **ńimiŋhīa-ndi-ʔ** təniʔīa bitir-ə-ʔ.*
          bigger.GEN   tale-LOC   NEG.AUX-CO-3PL   so   drink-EP-CNG
          'In the [real] big stories, people (they) usually don't drink.'
          [KNT_940903_KehyLuu.204]

      b. *ńükü əmiʔīa   miini-ə-ʔ   əmə-ʔ   taa-ʔ*
          child in.this.way here-ADJ-PL this-PL domestic.reindeer-3PL
          *əmiʔīa   muuntəmuŋhūa-ndu-ʔ   **ŋanaʔsan-ə-ʔ***
          in.this.way   guard-CO-3PL   person-EP-PL
          'My son, these are reindeers and that's how people guard them.'
          [KES_080721_Disease_flkd.016]

      c. *tə-tə   kuńiʔīa mej-čə-ŋi-ndəʔ   ŋanasan-ə-ʔ.*
          well-well   how   do-FUT-INTER-3PL.REFL   person-EP-PL
          'Well, what can people do about it.'
          [ChND_080719_Chunanchar_flkd.014]

---

19 Non-finite converbs have tense distinctions in Nganasan. They distinguish present, past and a periphrastic future tenses (see Wagner-Nagy 2019: 270–273).
20 This sentence was elicited from a native speaker [KNT].

d. *mənə   nanu-nə   ďaŋguj-śütə   ńaaďətə-tuə*
    I.GEN   with-POSS.1SG be.absent-FUT.3SG   speak.Nganasan-PTCP.PRS
    'In my opinion, one will not speak Nganasan.'
    [KSM_080804_Language_nar.006]

As we have seen, Nganasan allows all types of null-subjects across all sentences and tenses, including 3sg that need not be bound by an antecedent in the main clause (20d). Furthermore, Nganasan exhibits exclusive generic readings identified by 3pl verbal suffixes (23a) and inclusive readings with overt lexical DPs 'ŋanasan-ə-ʔ' (man/person-EP-PL 'people') (23b-c). The sentence in (23d) might be interpreted as an example with a dropped 3sg (identified merely by 3sg verbal suffixes) showing a generic reading.

4 Summary and conclusion

On the basis of the examples extracted from the corpora it has been shown that Selkup and Nganasan, at first glance, look like consistent NSLs. Quasi-null subjects (expletive null subjects) with weather verbs (see (12a-d) for Selkup and (21a-b) for Nganasan) cannot be overt. Furthermore, in both these languages, referential null subjects for all persons (1, 2 ,3) and number (SG, PL), across all tenses and sentence types, are attested (see (9a-e) for Selkup and (20a-d) for Nganasan). Embedded referential subjects do not need to be bound by higher arguments in main clauses (see (9d) for Selkup and (20d) for Nganasan), and the few examples with overt pronouns that are attested in the corpora show disjoint reference for the embedded null subjects, e.g. (7e) for Selkup. The data presented here seem to indicate that both Selkup and Nganasan are consistent NSLs. However, the analysis of the generic sentences serves to obscure the otherwise transparent picture. Both the languages exhibit different impersonal constructions with different generic readings. Typically, in a consistent NSL a 3SG null subject is associated with a referential [3SG] null-subject, whereas the generic reading is derived by using a special clitic (e.g. Italian or Spanish) or a verb from (e.g. Greek) (cf. D'Alessandro & Alexiadou 2002; Roberts & Holmberg 2010). This is not the case in Selkup and Nganasan, as they allow the generic reading with [3SG] null-subjects (see (16c) for Selkup and (23d) for Nganasan). All in all, we find three types of constructions with different readings. It seems that the inclusive reading is derived by a lexical DP (in Selkup it is the DP *qu-la* 'people' and in Nganasan the DP *ŋanasan-ə-ʔ* 'people'), and the generic reading calls for 3sg or 3pl null subjects. Especially, the fact that the exclusive generic reading in impersonal construc-

tions is derived by a 3sg null-subject, and not by an overt element, shows that the typology proposed by Roberts & Holmberg (2010) should rather be understood as a continuum.

References

Alexiadou, Artemis & Elena Anagnostopoulou. 2007. The subject-in-situ generalization revisited. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Proceedings of the Workshop on Interfaces + Recursion = Language*. Berlin: Mouton de Gruyter.

Borer, Hagit. 2001. Morphology and syntax. In Andrew Spencer & Arnold Zwicky (eds.), *The Handbook of Morphology*, 151–190. London: Blackwell Publishing.

Brykina, Maria, Svetlana Orlova & Beáta Wagner-Nagy. 2018a. INEL Selkup Corpus. Version 0.1. In Beáta Wagner-Nagy, Alexandr Arkhipov, Anne Ferger, Daniel Jettka & Timm Lehmberg (eds.), *The INEL Corpora of Indigenous Northern Eurasian languages*. Archived in Hamburger Zentrum für Sprachkorpora. http://hdl.handle.net/11022/0000-0007-CAE5-3, accessed on 2021-02-16.

Brykina, Maria, Valentin Gusev, Sándor Szeverényi & Beáta Wagner-Nagy. 2018b. *Nganasan Spoken Language Corpus (NSLC)*. Archived in Hamburger Zentrum für Sprachkorpora. Version 0.2. http://hdl.handle.net/11022/0000-0007-C6F2-8, accessed on 2021-02-16.

Budzisch, Josefina, Anja Harder & Beáta Wagner-Nagy. 2019. *Selkup Language Corpus (SLC)*. Archived in Hamburger Zentrum für Sprachkoropra. Version 1.0.0. http://hdl.handle.net/11022/0000-0007-D009-4, accessed on 2021-02-16.

Cabredo-Hofherr, Patricia 2003. Arbitrary readings of third person plural pronominals. In Matthias Weisgerber (ed.), *Proceedings of the Conference "sub7-Sinn und Bedeutung"*, 81–94. University of Konstanz.

Cinque, Gugliermo. 1988. On *si* constructions and the theory of ARB. *Linguistic Inquiry* 19. 521–581.

D'Alessandro, Roberta & Artemis Alexiadou. 2002. Inclusive and exclusive impersonal pronouns: a feature-geometrical analysis. *Rivista di Grammatica Generativa* 27. 31–44.

Dalmi, Gréte. 2014a. All-in-one: Generic inclusive null subjects in Hungarian. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Fomral Linguistics*, 115–123. Somerville, MA: Cascadilla Proceedings Project.

Dalmi, Gréte. 2014b. The feature geometry of generic inclusive null DPs in Hungarian. In Anna Bondaruk, Gréte Dalmi & Alexander Grosu (eds.), *Advances in the Syntax of DPs: Structure, Agreement, and Case*, 165–192. Amsterdam: Benjamins.

Egerland, Verner. 2003. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax* 71. 75–102.

Fischer, Susann. 2012. Subjects, subject positions and word-order: Old Romance vs. Old Germanic. In Jörg Fleischer & Horst Simon (eds.), *Comparing Diachronies*, 119–141. Berlin: Mouton De Gruyter.

Fischer, Susann. 2016. Existentials vs. unaccusatives: The definite-restriction in Romance. In Susann Fischer, Tanja Kupisch & Esther Rinke (eds.), *Definiteness Effects. Bilingual, Typological and Diachronic Variation*, 301–331. Cambridge: Cambridge Scholars Press.

Holmberg, Anders. 2005. Is there little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. The null generic subject pronoun in in Finnish: A case of incorporation in T. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 200–230. Cambridge: Cambridge University Press.

Holmberg, Anders, & Michelle Sheehan. 2010. Control into finite clauses in partial null subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 125–152. Cambridge: Cambridge University Press.

Manzini, M. Rita & Leonardo M. Savoia. 1997. Null subjects without *pro*. *UCL Working Papers in Linguistics* 9. 1–12.

Perlmutter, David. 1971. *Deep and Surface Constraints in Syntax.* Nex York: Holt Rinehart and Winston.

Rizzi, Luigi. 1982. *Issues in Italian Syntax.* Cambridge, MA: MIT press.

Rizzi, Luigi. 1986. On the status of subject clitics in Romance. In Osvaldo Jaeggli & C. Silva-Corvalán (eds.), *Studies in Romance Linguistics*, 391–420. Dordrecht: Foris.

Roberts, Ian & Anders Holmberg. 2010. Introduction. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–58. Cambridge: Cambridge University Press.

Wagner-Nagy, Beáta 2019. *A Grammar of Nganasan*. Leiden – Boston: Brill.

Piotr Cegłowski
13 Concluding remarks

It's been exactly five decades since Perlmutter (1971) brought the concept of *pro*-drop into the limelight of generative studies. These fifty years were devoted to intense comparative research seeking to explore numerous fascinating, previously undiscovered, mutually related cross-linguistic facts concerning the issue of dropped, i.e. phonologically null subjects. Given all this time and effort, a great amount of linguistic data followed by observations, analyses and axioms, running virtually into hundreds, one may wonder whether there's still room for new discoveries and ideas in the field. Is there anything else we can still learn about the nature of null subjects and the characteristics of Null Subject Languages?

For a scientist, whether a linguist, a biologist, or a molecular physicist, such dilemmas can be trivially rejected. As scholars, we are all acutely aware of the fact that there is no such thing as a foolproof hypothesis or a perfect theory. In other words, a theory is valid as long as some new theories and results arise that falsify it. A picture emerging from the empirical data is transparent and coherent until some new data serve to obscure it. A non-falsifiable hypothesis is, by definition, not a good hypothesis. This recognition encourages us to believe that the present volume will shed some more light on the subject matter of the null subject phenomenon.

On a less trivial note, the contributions collected here advance in several paths which have been amply delineated in Chapter 1 by Jacek Witkoś. For readers with no comprehensive knowledge of the phenomenon, the chapter is an 'orientation tour' that constitutes a firm point of departure to the chapters to follow.

The multi-faceted, cross-linguistic analysis of languages from the Slavic Baltic, Finno-Ugric and Samoyedic language families has revealed a great deal of interesting findings concerning the nature of the typological distinction, and the notion of NSLs. In particular, at least a couple of Null Subject Languages presented in this volume do not seem to uniformly follow the criteria set for one particular NSL-type. Thus, in her elaborate analysis of Bulgarian in Chapter 6, Dobrinka Genevska-Hanke has discussed empirical data supporting the claim that a more comprehensive approach to the traditional criteria for a consistent NSL (cf. Beletti 2009; Roberts 2014) – one that allows for microvariation – is required. While, according to Genevska-Hanke, Bulgarian is a consistent NSL, it does not easily fit in with a clear-cut classification.

An interesting aspect of the null subject phenomenon contemplated in this volume has been the role of the null/overt dichotomy (for the status φP and DP, see Holmberg 2005, 2010; see also Cardinaletti & Starke 1999; Déchaine & Wiltschko

2002). In the course of his analysis of root clauses in East Slavic languages (Belarusian, Russian and Ukrainian) Egor Tsedryk in Chapter 2 adopted Ackema & Neeleman's (2018) framework to analyse the correlation of person features (PROX, DIST), responsible for the distribution of a (broad spectrum of) subject pronouns. Tsedryk also attributes the partial *pro*-drop phenomenon to the visibility condition regulating minimal links within the clausal domain (Frascarelli 2018). In turn, the aim of Lida Veselovská's analysis of Czech null subjects in Chapter 5 was to correlatate the two types of φ-features (unique index agreement and recursive concord agreement) and their role in the valuation of T. Additionally, Veselovská considers another way to value the unvalued features of T (an alternative scenario to the D-to-T pronominal clitic incorporation), i.e. the valuation by the Topic head.

On a diachronic note, Nerea Madariaga (Chapter 3), departing from the assumption that null subjects are φPs (Holmberg 2005), examined the transition from Old Russian, classified as a "well-behaved" consistent NSL, to Modern Russian, a partial NSL. Madariaga's analysis of the switch from the former (Grammar 1) to the latter (Grammar 2) concentrates on the change that affected the featural composition of T, i.e. the loss of the unvalued D-feature that eventually served to cancel the T's ability to mediate in referential index transmission from a null Topic / operator at CP to the null subject.

The role of the left periphery in the licensing and interpretation of null subjects, i.e. the Aboutness-Shift Topic (see Frascarelli 2007, 2018; Holmberg et al. 2009; Holmberg 2010) and the Operator Field (Bródy 1995) was the main research question of Marta Ruda's critical evaluation of left-peripheral linking Polish (Chapter 4) and Gréte Dalmi's proposal for the licensing mechanisms of lexical and null pronominal arguments in Hungarian (Chapter 9). Starting with the former, Marta Ruda scrutinized both the conceptual nature and the empirical application of the Topic Criterion as proposed by Frascarelli & Jiménez-Fernández (2019). Applying "the lens of explanatory adequacy" she hints at the unexplained nature of both the linking relation itself and the fact that both the types of topics (Aboutness-Shift, Given) can be null. Focusing mainly on Polish, she argues convincingly that the language poses a non-trivial challenge to the left-peripheral linking mechanism proposed by Frascarelli (2007). As for the latter, Gréte Dalmi investigated the licensing conditions for lexical and null pronominals with the individual and generic reference interpretations in Hungarian. In her model, it is sufficient for lexical pronominal arguments with the individual reference interpretation appearing in dependent clauses to be syntactically licensed by an antecedent. She put forward a dual licensing mechanism for their null counterparts, which must be licensed both syntactically, via fearure agreement, and semantically, by being in the scope of an operator in the Operated Field (Bródy 1995). In

line with Bianchi (2006), Dalmi further proposed that generic lexical and null arguments are semantically licensed by GN sitting in SAPP, which she took to be the leftmost position of the C-domain. In her classification, the mixed NSL-properties of Hungarian place it somewhere between partial and radical NSLs (cf. Roberts & Holmberg 2010).

Apart from the mainstream minimalist research, the volume offered various alternative views of the *pro*-drop phenomenon. Thus, in his analysis of null subjects in Finnish finite sentences in Chapter 8, Urpo Nikanne has investigated root finite clauses, conducting the study within the micro-modular theory of Conceptual Semantics (Nikanne 2018). As Nikanne himself points out, the theory departs from the syntax-centered theories of grammar that stress the pivotal role of syntactic feature checking in the derivation of the sentence. Null subjects are interpreted by way of a three-step procedure involving the recognition of the lexical items, the comparison of the syntactic structure and the lexical conceptual structure, as well as identifying the linking rules between the two structures.

The analysis of Mari by Jeremy Bradley and Johannes Hirvonen in Chapter 10 has focused on the distribution of overt and null pronominal subjects across different syntactic contexts. Bradley and Hirvonen conclude that Mari generally follows the traditional criteria of consistent NSLs in the sense of Holmberg (2010) and Roberts & Holmberg (2010), among others, yet it evades absolute generalizations.

The assumption that null subjects are φPs also underlies Mikael Vinka's analysis of the third person null subjects in South Saami in Chapter 11. Specifically, Vinka has examined the behaviour of third person null subjects in both indicative clauses and unselected subjunctives, as well as selected subjunctive complements. By looking at the specific syntactic contexts, he concludes that while pro in the former is indeed φP, there are reasonable grounds to assume that the null subject in the latter is a pronoun that cliticizes to the Mood head and subsequently undergoes PF-deletion. On Vinka's account, the obligatory coreference of the null subject of the subjunctive complement with either argument in the higher clause may be attributed to the participant feature that this pronoun carries, i.e. [AUTH/ADDR] (cf. Landau 2015).

Following an extensive overview of the corpus data from Selkup and Nganasan (Chapter 12), Susann Fischer and Beáta Wagner-Nagy have hinted at an intriguing discrepancy, i.e. the fact that both these languages (otherwise exerting all the characteristics of a consistent NSL) allow 3sg null subjects with the generic reading. Based on their analysis of generic exclusive readings in impersonal constructions, Fischer & Wagner-Nagy show conclusively that Selkup and Nganasan give every reason to consider the typology put forward by Roberts & Holmberg (2010) in terms of a continuum, rather than a clear-cut, categorical classification.

To sum up, there are good reasons for which this intellectual journey is worth taking. First and foremost, the discussions, analyses and postulates put forward in the respective chapters add numerous new pieces to the otherwise conceptually and empirically complex puzzle of *pro*-drop. At this point, let me make a humble disclaimer, even though it may seem obvious to everyone who has just read this section; in particular, the concluding summary of the content sketched above represents but one way to do justice to the immense, multi-dimensional structure of the phenomenon that emerges from the specific contributions collected here. Second, the linguistic material presented, whether gained by introspection, elicited from language users by experimental means or obtained as a result of an in-depth corpus study, may serve in the future as a useful point of reference for scholars representing various approaches and conducting their research in different fields of linguistics. Third, by attracting attention to less studied languages, the book helps fill in certain gaps in the typological classification of NSLs, thereby adding substance to the specific criteria and distinctions put forward in the literature. Last, but not least, its didactic value as a comprehensive resource for teachers and students of advanced linguistics courses should also be acknowledged.

References

Ackema, Peter & Ad Neeleman. 2018. *Features of Person: From the Inventory of Persons to their Morphological Realization*. Cambridge, MA: MIT Press.

Belletti, Adriana. 2009. *Structures and Strategies*. London: Routledge.

Bianchi, Valentina. 2006. The syntax of personal arguments. *Lingua* 116. 2023–2067.

Biberauer, Theresa. 2018. *Pro*-drop and emergent parameter hierarchies. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar*. *A Synchronic and Diachronic Approach*, 94–134. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge Handbook of Historical Syntax*, 134–162. Cambridge: Cambridge University Press.

Bródy, Mihály. 1990. Remarks on the order of elements in the Hungarian focus field. In István Kenesei (ed.), *Approaches to Hungarian* 3. 95–123.

Cardinaletti Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk Van Riemsdijk (ed.), *Clitics in the Languages of Europe*, 145–233. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: The MIT Press.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33. 409–442.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. *Natural Language and Linguistics Theory* 25. 691–734.

Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial null-subject languages. A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null Subjects in Generative Grammar. A Synchronic and Diachronic Approach,* 211–239. Oxford: Oxford University Press.

Frascarelli, Mara & Ángel L. Jiménez-Fernández. 2019. Understanding partiality in *pro*-drop languages: An information-structure approach. *Syntax* 22 (2–3). 162–198.

Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97.

Jaeggli, Osvaldo & Kenneth Safir. 1989. The null subject parameter and parametric theory. In Osvaldo Jaeggli & Kenneth Safir (eds.), *The Null Subject Parameter*, 1–44. Dordrecht: Kluwer.

Landau, Idan. 2015. *A Two-tiered Theory of Control*. Cambridge, Mass.: MIT Press.

Nikanne, Urpo 2018. *Conceptual Semantics: A Micro-modular Approach*. Amsterdam & Philadelphia: John Benjamins.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Cambridge, Mass.: The MIT Press.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–557.

Roberts, Ian. 2014. Taraldsen's generalization and language change: two ways to lose null subjects. In Peter Svenonius (ed.), *Functional Structure from Tip to Toe. The Cartography of Syntactic Structures, Vol. 9*, 115–147. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Parameters in minimalist theory. In Teresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan, *Parametric Variation: Null Subjects in Minimalist Theory*, 1–57. Cambridge: Cambridge University Press.

Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality-theoretic Analysis*. PhD dissertation. Rutgers University, New Brunswick, NJ.

Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense and speech features. *Italian Journal of Linguistics* 16. 219–251. (Special issue edited by Valentina Bianchi & Kenneth Safir).

Index