Structuring Variation in Romance Linguistics and Beyond

In honour of Leonardo M. Savoia

Edited by

Mirko Grimaldi

Rosangela Lai

Ludovico Franco

Benedetta Baldi

John Benjamins Publishing Company

Structuring Variation in Romance Linguistics and Beyond

Linguistik Aktuell/Linguistics Today (LA)

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/la

Founding Editor

Werner Abraham
Universität Wien / Ludwig Maximilian Universität München

General Editors

Werner Abraham
Universität Wien /
Ludwig Maximilian Universität München

Elly van Gelderen
Arizona State University

Advisory Editorial Board

Josef Bayer
University of Konstanz

Cedric Boeckx
ICREA/UB

Guglielmo Cinque
University of Venice

Liliane Haegeman
University of Ghent

Hubert Haider
University of Salzburg

Terje Lohndal
Norwegian University of Science
and Technology

Christer Platzack
University of Lund

Ian Roberts
Cambridge University

Lisa deMena Travis
McGill University

Sten Vikner
University of Aarhus

C. Jan-Wouter Zwart
University of Groningen

Volume 252

Structuring Variation in Romance Linguistics and Beyond
In honour of Leonardo M. Savoia
Edited by Mirko Grimaldi, Rosangela Lai, Ludovico Franco and Benedetta Baldi

Structuring Variation in Romance Linguistics and Beyond

In honour of Leonardo M. Savoia

Edited by

Mirko Grimaldi
University of Salento

Rosangela Lai
University of Florence

Ludovico Franco
Nova University of Lisbon

Benedetta Baldi
University of Florence

∞™ The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI z39.48-1984.

Table of contents

CHAPTER 1

Introduction: Structuring thought, externalizing structure
Variation and universals

M. Rita Manzini
University of Florence

The articles collected in this volume bear witness to the important role played by studies on language variation, specifically on Romance microvariation, in the development of formal approaches to language. Leonardo M. Savoia is a central figure in this endeavor, in the twin fields of phonology and morphosyntax. In this introductory chapter, I will briefly put research into (micro)variation in Romance and other language families in the context of the recent theoretical debate in generative grammar. Despite Savoia's unrivalled expertise in field work and the prodigious number of data we owe to him, especially on Italian and Albanian varieties, it is only within a theoretical setting that his work is truly appreciated. I will conclude each section by illustrating how the themes explored therein resonate in the several contributions to the volume.

1.   Is phonology different?

Language externalizes thought, integrating cognitive processes as different as perception/articulation and categorization/inference into a single behavior. For Hauser et al. (2002) it is the Faculty of Language in the Narrow sense (FLN), essentially recursive syntactic computation, that represents the crucial innovation of human language. FLN recruits around itself preexisting conceptual-intentional systems and sensory-motor systems, yielding the Faculty of Language in the Broad sense (FLB). For Berwick and Chomsky (2011), Chomsky et al. (2018) externalization (EXT) is the locus of variation, hence essentially morphonology and the lexicon, face to the universal nature of semantics (SEM) and syntax.

This simple picture leaves a number of fundamental issues open. The most obvious one, concerns the relation of phonology to syntax. The history of generative grammar leaves little doubt that phonology is a computational system. An early classical result is the cross-domain application of key computational

notions of locality, originally defined by the notion of cycle (Chomsky & Halle 1968; Chomsky 1973) currently by that of phase (Chomsky 2001; D'Alessandro & Scheer 2013; Newell & Piggott 2014). These are seen to constrain both syntactic operations (movement) and phonological ones.

The phonological model that is most explicit in advocating a parallelism between standard syntactic models (Chomsky 1981, 1995) and phonology is government phonology, starting with the seminal work of Kaye et al. (1985, 1990) Most of Savoia's phonological work is explicitly couched within the framework of government phonology, and in opposition (explicitly or implicitly) to richer frameworks, such as the long-dominant Optimality Theory framework. Having co-authored one of the most widely cited works on consonant weakening in the Italian and Romance domains (Giannelli & Savoia 1978, 1979–80), Savoia's work on the phonology of Italian and Albanian varieties then came to focus especially on the fine variation defined by the interaction between metrical and syllabic structure and vocalic realization with specific attention to harmonic processes, i.e. so-called metaphony (Savoia 1990a, 1994a, 1997a, 2005; Savoia & Maiden 1997; Savoia & Marotta 1994). This work, using Italian (and Albanian) microvariation to pry some of the most classical problems in theoretical phonology, culminates in two recent monographs (Savoia 2014, 2015a) and in a spate of very recent articles (Savoia 2016; Savoia & Baldi 2016, 2017).

Among the results that more strongly characterize Savoia's phonological work is the treatment of metaphony within the government phonology framework. Metaphony is a vowel harmony phenomenon traditionally described as triggered by a final high vowel, though the latter is very often only etymologically reconstructed, determining the raising or tensing or diphthongization of the tonic vowel of the word. There are several aspects of this traditional description that are incompatible with a restrictive framework such as government phonology. For instance, harmony processes are expected to be triggered by tonic vowels. Savoia proposes that indeed this is the case in Italian varieties. The tonic nucleus must have positively specified high properties in order to be able to license the same properties in the final nucleus. Metaphonic processes are prevalent in South Italian varieties though not limited to them. The extremely fine microvariation found in these varieties allows Savoia to explore in great detail the varying conditions under which harmonic processes hold.

The adoption of the government phonology framework is motivated by phonology-internal reasons, but corresponds also to a clear endorsement of the continuity between phonological and syntactic modules. We single out a couple of themes that are especially important in this respect.

We briefly mentioned formal universals bridging phonology and syntax, specifically locality. Within the government phonology tradition we may add at least

the (controversial) Projection Principle, essentially to the effect that the derivation cannot add information not projected from terminals – nor undo structure projected from them. What we want to concentrate on next is substantive universals, i.e. the primitives on which (syntactic or phonological) computation is built. Needless to say, two entirely different sets of properties, namely articulatory/acoustic and conceptual ones are involved. Nevertheless the question can meaningfully be asked whether the primitives are binary features (Chomsky and Halle 1968), privative features organized by feature geometries – or simply primes, i.e. positively specified self-standing contents (Kaye et al. 1985). If there are no phonological feature, but phonological elements such as the vowel defining contents |A|, |I|, |U| – then phonological primitives are a lot more like syntactic primitives, namely interpretable absolute categories such as D or N. The choice of primes as phonological primitives predicts that morphology will also be built on categorial contents, not binary or privative features.

Independently of how homogeneous (or heterogeneous) the computational/representational systems of phonology and syntax are, their interface represents an issue. The issue arises precisely because, notwithstanding formal similarities, the two systems are run on different primitives.

Natural languages are rife with phonological processes that read syntactico-semantic structures, namely so-called phonosyntactic rules. Two approaches are in principle possible. In a grammatical architecture consisting of encapsulated modules, the existence of operations in a given module sensitive to properties and relations pertaining to a different module is not expected. To the extent that it is found, it may be addressed by precompiling information readable by level x into level y. The questions to be asked is whether this kind of precompiling is necessary or advantageous – if neither is the case, then we should conclude in favor of the legibility of syntactico-semantic structures by externalization (and conversely).

This issue is discussed by Savoia in relation to harmonic processes once again, namely so-called /u/ propagation in some South Italian varieties (Savoia 1987a; Rizzi & Savoia 1993; Manzini & Savoia 2016a), which takes place both word-internally and in syntax according to finely varying conditions. For Rizzi and Savoia (1993), accounting for microvariation requires abstract government, mutual government and agreement relations to be visible in the phonology. For Manzini and Savoia (2016a) externalization can access any syntactico-semantic relation. For instance, some South Italian dialects have /u/ harmony only in the clitic–V configuration and in the D-N configuration, corresponding to saturation of the N/V predicate by the D/clitic argument (Higginbotham 1985). In other varieties, harmony is extended to A-N configurations, corresponding to Higginbotham's theta-identification relation.

The articles collected under the title of 'Sound Patterns and Syntactic Structures' reflect some of the phonological themes illustrated in this section. Passino's contribution addresses a core prediction of phonological models, namely that syllable structure is only determined by the melodic composition of segments. She considers evidence from the several Italian varieties showing that different syllabic parses exist for the same cluster in a given language and that unpredicted syllabic parses are documented cross-linguistically, dismissing facile approaches such as denying the centrality of the data or adding ad hoc theoretical machinery to draw them in. Of particular relevance for the themes debated above is the article by van Oostendorp and D'Alessandro. They point out that within elements theory, the raising of vowels that characterizes metaphony is difficult to express to the extent that it involves unrelated |I| and |U| primes; therefore it is best expressed as the subtraction of an |A| element. D'Alessandro and van Oostendorp (2016a) in turn propose that this is best expressed as the linking of the |A| element to an empty suffix that represents in fact the plural morphology of the language; recall that many South Italian varieties have word final schwa so that metaphony is in fact the externalization of plural in the language. The article in this volume provides a theoretical framework that allows them to refine the previous proposal, specifically allows them to avoid the stipulation that the suffix attracting |A| must be empty.

The contribution of Ledgeway is a study of the classical phonosyntactic phenomenon of *raddoppiamento* based here on the data of the Calabrian variety of Cosenza, where the phenomenon is triggered by a limited repertory of items (as is not infrequently the case). Ledgeway, in line with recent developments hinted at above, argued that the only locality constraint active on *raddoppiamento* is that the items involved be internal to the same phase. Floricic and Molinu also consider an issue at the interface between syntax and externalization, namely phonologically constrained truncation as an externalization of the syntactically defined category of vocatives. Specifically, in Sardinian varieties the Vocative is formed deleting all the material following the stressed syllable; this process is studied in great detail also in relation to the proposal of D'Alessandro and van Oostendorp (2016b).

2.   The unification of morphology and syntax

In the previous section we have avoided reference to morphology, for the very fundamental reason that its exact status in the architecture of grammar is controversial. Hallean morphology, codified as Distributed Morphology (DM) by Halle and Marantz (1993), implies that the listed elements of the mental lexicon are roots and morphemes and that the computational operation of Merge covers not

only constituent structure but also word formation ('syntax all the way down'). This approach is by no means unchallenged within formal models of grammar (Blevins 2006); nevertheless, following Manzini and Savoia (2005, 2007, 2011a) among many others, we adopt it here without further discussion.

The aspect of the standard DM model that is strongly challenged in the work of Manzini and Savoia is the existence of a Morphological Structure component (in the sense of Halle & Marantz 1993) hosting powerful morphological rules and acting as a buffer between an abstract syntax and its externalization by the lexicon (or rather the Vocabulary Items), under Late Insertion. In fact, the issue goes deeper than this. The Late Insertion postulate of DM (required by the opacization operations available within the model) becomes the passepartout for an abstract syntax whose relation to the externalized string is potentially arbitrary. While largely invariant syntactic structures isomorphic to interpretation have been held advantageous especially in the cartographic tradition (cf. the 'syntacticization' of semantics proposed by Cinque & Rizzi 2010), the advantage in the Uniformity of syntax and semantics (in the sense of Culicover & Jackendoff 2005) is paid for in terms of a realizational conception of the lexicon and morphology, which in principle allows for a completely arbitrary matching of morphophonology and syntax.

In the model of Manzini and Savoia, based on standard minimalism (Chomsky 1995), the lexicon, understood as the pairing of phonological and conceptual content to form lexical items, is prior to syntax. Syntax therefore projects lexical items to syntactic structures. Crucially, given a syntax projected from an abstract lexicon, it becomes impossible to conceive post-syntactic operations which manipulate abstract terminals. Specifically DM theorists assume that there are at least three such operations, namely Fusion (two abstract terminals are fused – effectively the morphological equivalent of an incorporation in the syntax); Fission (an abstract terminal is split apart and recombined with other terminal – effectively the morphological counterpart of an excorporation/movement operation in the syntax); and Impoverishment (the deletion of abstract terminals or features). Impoverishment contributes a way to handle syncretisms; once abstract terminals are suitably impoverished (specifications deleted), terminals which are distinct in the syntax, end up identified in Morphological Structure and hence paired with the same sensory-motor labels, yielding syncretism. However, the hypothesis that syncretisms require no Impoverishment, hence no morphological readjustment, has occasionally been pursued by the syntactic literature, not only by Manzini and Savoia, but also for instance by Kayne (2010). The same terms of the debate are recognizable for phenomena treated in terms of Fission or Fusion by DM; thus infixation (for instance of clitics, i.e. mesoclisis) does not require special morphological handling but can be treated in terms of ordinary syntactic movement (Manzini & Savoia 2011c).

Issues relating to nominal inflection have been the prevalent empirical focus of morphosyntactic research by Savoia and colleagues. The most complex such systems in most Romance languages are provided by clitics and personal pronouns – whence the particular attention reserved to these elements both in the works quoted and in many related articles (Manzini & Savoia 2002a, 2004a, 2010a, 2010b, 2011d, 2011c, 2017c; Savoia 2007a; Savoia & Manzini 2010a, 2010c, 2015). At the same time, the hypotheses initially arrived at on the basis of the Romance pronominal data were systematically controlled on languages with complex nominal declensions, including case, and specifically on Albanian varieties, again on the basis of rich variation corpora collected directly by Savoia (Manzini & Savoia 2011b, 2012e, 2014c; Savoia & Manzini 2010c), and on Eastern Romance varieties, including the minority language Aromanian (Manzini & Savoia 2017b). In turn, the nominal inflection systems of languages like Italian, precisely because of their apparently elementary nature, hide a number of insidious and difficult issues concerning the nature of fundamental notions of number and gender, or more generally of nominal class (Manzini & Savoia 2017a, 2017b). The study of evaluative morphology allows Savoia to show the labile nature of another traditional distinction, that between inflectional and derivational morphology (Savoia et al. 2017a, b).

Let us consider some of the theoretical issues raised by the various phenomena just evoked. As we saw, according to DM, syntax operates on abstract features roughly corresponding to the descriptive categories of traditional grammar. Opacization operations, which blur the syntactic full feature specification (by unrecoverable deletion, i.e. Impoverishment, by Fusion, by Fission) yield syncretisms as well as portmanteau morphemes or discontinuous morphemes. The position argued for by Manzini and Savoia is that syncretisms correspond to natural classes (or to say it with Kayne, that syncretic morphemes are simply ambiguous). Manzini and Savoia further argue that this is deemed to be too strong a position face to empirical evidence simply because the discussion assumes the traditional repertory of categories and features as given. Providing a projectionist, unified model of morphosyntax implies reforming morphosyntactic categories/features.

Consider for instance the syncretism of plural number and oblique case seen in Latin *-i*, Latin *-s*; the former is continued by the pronominal clitics of Italian where the two allomorphs *gl-i* and *l-i* instantiate both dative and plural (masculine). Halle and Vaux (1998), cf. Calabrese (1998, 2008), essentially renounce giving an analysis; *-s* is a morpheme with no content (the default of the system) and there are two separate *-i* morphemes. In consonance with recent literature on number (Borer 2005) Manzini and Savoia take it that plural is effected by a partition operator, individuating a subset in the denotation of the predicate. They propose that the same operator is naturally construed as charactering one of the

fundamental components of the oblique, i.e. the partitive, and by extension all cases introducing a possession relation (genitive proper, dative), construed as a loose part-whole relation ('zonal inclusion' for Belvin & den Dikken 1997). In this latter instance the inclusion elementary predicate/operator applies to phrasal syntax, introducing a relation between two phrasal constituents. For instance *John's nose* says that the nose is part of/possessed by John.

Independently of whether the proposed unification of number and case is correct, the point to be appreciated here is that the traditional oblique – plural syncretism depends on the same subset/inclusion content interacting with syntactic computation. In short, traditional categories capture superficial generalizations that are useful for descriptive purposes, but do not correspond to the real computational organization of language. The use of realistic categories and features reveals that in some, or many, instances what appear to be quirks of the externalization process, in reality correspond to deeper generalizations within the syntactico-semantic component. In a similar vein traditional notions of gender and number reveal inadequate; rather the morphological structure of N is better accounted for by acceding to the idea that, as Nominal classes in Bantu or classifiers in Chinese partition roots between a few fundamental natural classes of objects, so does gender (see especially Déchaine et al. 2014), and in fact number (see the discussion of divisibility above).

On the other hand, it is to be expected that the interfacing of two computational systems based on two different sets of primitives cannot be entirely governed by isomorphic mapping. Functional optimization conditions, corresponding the 'third factor' of Chomsky (2005), are expected to play a role. In recent work devoted to clitic allomorphies and reordering in Romance, governed by so-called enclisis/proclisis alternations, Manzini and Savoia (2017d) make a plea against "the idea that allomorphies and other devices traditionally considered to be morphological simply disrupt the underlying regularity of syntactico-semantic structure. On the contrary they are seen to contribute to the externalization of complex CI information". In essence, segmentally and accentually richer clitic forms are found in the syntactic scope of modals (imperatives, negation). Manzini and Savoia's "read this requirement as a device to optimize visibility at the SM interface of two key operators (D and C/Neg) interacting at the CI interface". In other words, PHON crucially cooperates to the optimum visibility of SEM – against conventional wisdom linking the SEM/PHON interface to opacization devices.

The contributions collected in the 'Clitics and pronouns' sections of this book are those that most directly reflect the issues raised in this section, sometimes adopting the highly distinctive line of analysis advocated by Savoia and colleagues and sometimes diverging from it. Bafile and Lai discuss the interaction of clitics and stress in enclisis/proclisis alternations on the basis of original Sardinian data.

Their detailed and careful phonological discussion allows them to conclude stress shift cannot be solely accounted for in terms of metrical requirements. In assessing weak pronouns theories, they explore hitherto unreported data, concerning the presence of words stress on proclitic groups in the relevant Sardinian varieties. In the article by Garzonio and Rossi, clitics are present as triggers of vocalic syncope/epenthesis in non-rizotonic verb forms (e.g. *le'zi* vs *al'zi* 'you(pl) read') in three closely related varieties of Northern Emilia. The alternation between epenthetic and non-epenthetic forms is interpreted as an allomorphy conditioned by the presence of the clitic.

The article by Roberts takes us in the domain of mutual exclusions between subject clitics (SCL) and object clitics (OCL), specifically of a range of phenomena which Roberts characterizes as OCL-for-SCL: "an auxiliary always requires a proclitic; if this is a subject clitic, the object clitic is enclitic to the past participle, but [if] the object clitic [is] proclitic on the auxiliary, [...] the subject clitic does not appear". For Manzini and Savoia (2005) these phenomena are just part of anti-identity conditions on clitic strings; Roberts argues for an analysis within the framework of Roberts (2010), integrated by a DM-like morpho-phonological component. The final contribution to be mentioned in this section is Grossman and Thornton's on the object pronouns of Hungarian. The singular is characterized by a person split whereby the accusative suffix *-et* may be reduplicated in 3P while on the contrary it may be omitted in 1/2P. As for the 1/2PL, they are built either from the *mi-, ti-* bases or from a *benn-* base, from **bel* 'internal part' followed by the locative *-n*; either base is followed by the 1/2PL possessive and by the accusative suffix. Grossman and Thornton are interested in overabundance, namely "the situation in which more than one inflected form is available to realize a single cell of an inflectional paradigm" and on the relative frequency of the forms. In a different perspective the data seem to connect with morphosyntactic phenomena such as the interaction of person splits and object marking (a form of DOM, see also Section 3).

Loporcaro, Romagnoli and Wild target Northern Sardinian varieties. Their object of study are number and gender inflections in nominal and pronominal system, where they bring out, among other facts, the very interesting masculine/feminine syncretism in sigmatic plurals (-ɔs in Sennorese, – *as* on Lurese Ds). They suggest an external interpretation of these alignments on Gallurese (-*i* gender-neutral plurals) in conditions of contact. Of special interest in the context of this volume is their conclusion that "as soon as the dialects of Italy are considered in-depth, typologically interesting facts emerge" – a matter to which we shall return in Section 3.

Two contributions that appear in the 'Micro- and macro-variation' section of the volume are best discussed in this connection. The first one, by Baldi and

Franco, addresses issues of nominal inflection. Baldi and Franco take as their starting point the internal structure of N proposed by Savoia et al. (2017a,b) and go on to provide an account of Northern Italian dialect data from the corpus of Manzini and Savoia (2005) within this framework. Their key idea is that the plural inflection -*i*, which is the only surviving nominal inflection apart from the feminine singular -*a*, is clearly endowed with semantic content – namely with the divisibility or part/whole content that underlies plurality. Finally in his contribution, Lorusso notes that the psycholinguistic literature yields a clear prominence of person over gender and number. However, though number is traditionally given primacy over gender, the psycholinguistic evidence is not clear-cut on this point. Specifically recent ERP studies have not found any electrophysiological effect confirming the different status of gender and number. Lorusso's conclusion is particularly relevant to the discussion in this section namely that gender and number are the traditional labels of what are in fact just different nominal classes.

3.   Syntactic parameters: Microvariation and macrocategories

In the research framework outlined in Section 1 (Chomsky & Berwick 2015; Chomsky et al. 2018), language is a biologically defined object, including universal and innate components, both conceptual and computational. At its core, linguistic variation is interesting to the extent that, no less than language universals, it appears to be not accidental to language, but rather written into its basic design. Specifically, conceptual and inferential components are presumably not subject to variation/change – nor are the computational components (syntactic operations, phonological operations). Rather, variation arises when these components interact (notably at externalization).

A major locus of variation is the interface connecting syntactico-semantic computation to phonological computation, and involving presumably functional optimization processes (see Section 2). The lexicon is the other natural locus of morphosyntactic variation, providing a double interface, between the conceptual system and syntactico-semantic computation on the one hand, and between syntax and externalization on the other hand. In the words of Berwick & Chomsky (2015: 83): "Why are there so many languages? The reason might be that the problem of externalization can be solved in many different and independent ways". In other words, "diversity of language results from the fact that the principles do not determine the answers to all questions about language, but leave some questions as open parameters".

This general conception does not in principle exclude that parametric space is structured, for instance by the hierarchical ordering of the values of a given

parameter or by the interaction between different parameters and/or their settings, such that a particular parameter or parameter value may influence another. Structured parametric spaces have been identified with a notion of macroparameter (Baker 2008) which is seen to be in opposition to work highlighting the microparameter as the minimal units of syntactic variation (Kayne 2010). Recent research seems to be characterized by an attempt to bring parameters theory in line with current minimalist theorizing. For instance, Biberauer et al. (2014) provide hierarchical schemas able to encode the best known extant parameters: the null subject parameter, the head-movement parameter, the (case) alignment parameter, the A'-movement parameter. Crucially, the order is not provided by the theory of grammar, but rather by functional principles. Yet the schemas are identically organized only at very general level; in the detail, each parametric template imposes its own organization – nor is it clear how the templates interact with one another.

Manzini and Savoia (2011a) take a much weaker approach, namely that "the parameters interacting with [Externalization] are … the categorical splits" for instance "speaker vs. hearer, 1st/2nd person vs. D". Thus even parameter schemas are epiphenomena. Underlying the latter are the categorial cuts (parameters) that are or are not externalized by the lexicon. Interactions between categorical cuts (parameters) in turn simply take the form "categorial split A is not defined for value 0/1 of … categorial split B". This weaker approach has the advantage of eliminating elements of rigidity present in all extrinsically ordered organization, thereby proving maximally suited to account for microvariation. The conception is microparametric in the precise definition by Kayne (2010), whereby a microparameter is an atomic parameter. At the same time, there is no question that such categorial splits as 1/2P vs. 3P – or better their interaction with the (invariant) computational system and externalization devices – also define major typological variation, as seen specifically in Person ergativity splits or other case alignments (e.g. DOM).

The research of Savoia and his colleagues in the domain of morphosyntactic variation is strongly characterized by an awareness of the theoretical stakes. We have already mentioned research in the empirical domains of pronominalization and nominal inflection. A natural extension of these studies are proposals on oblique case including: DOM as obliquization, especially in Albanian, the connections with the Person Case Constraint (Manzini & Savoia 2012e, 2014c), ergative alignments, with special reference to Punjabi (Manzini et al. 2015), the embedding of adnominal oblique modifiers under so-called Linkers, again in Albanian but also in Kurdish (Manzini & Savoia 2011a, 2014b; Franco et al. 2015a). The long-standing interest of Savoia in Arbëresh causatives may also be mentioned in this connection (Savoia 1989a, 1989b; Brandi & Savoia 1990; Manzini & Savoia 2007), since it is characterized by a case realignment of the Romance type (with

the causee as a dative or *by*-phrase) though the embedded sentential complement is finite, as required by infinitive-less Albanian varieties.

Complementation and complementizers have an equally important role in the syntactic work of Savoia and his collaborators. Comparison between Romance and Albanian or Balkan languages in general is traditionally centered on the absence of non-finite complementation, but study of dialectal variation reveals that Aromanian has so-called long infinitives, Geg Albanian has the so-called *paskajore* (and even Romanian has so-called supines). The various facets of finite and non-finite complementation are studied in close connection with the issue of sentential introducers. Based on their lexicalization by wh- operators and, by prepositions in the languages under consideration, but also on their syntax and semantics, Savoia and coauthors connect complementizers to the nominal system. Thus the embedding of sentences is not possible unless they are turned into relatives of sort (via wh- complementizer) or they are embedded under oblique case prepositions or possibly other strategies are employed cross linguistically (Manzini & Savoia 2003a, 2007, 2011a).

Several of the issues briefly discussed in Section 2 because of their interest for the morphological component of externalization have equally relevant syntactic implications. Thus recall that the functional hierarchies of cartographic approaches are deemed to be empirically inadequate, as is the complex array of movement rules necessary to support them (see also Chomsky et al. 2018). This translates into proposals that the cartographic Neg, or Adv categories be abandoned in favour of recognizing that negations are just Qs or bare Ns (minimizers, cf. Manzini & Savoia 2002c, 2011a, 2012d). Wh- in situ in Romance is shown not to undergo a remnant movement derivation (Manzini & Savoia 2011e). Particularly complex from the point of view of the double interface between syntax and externalization and between syntax and interpretation is middle-passive voice (Manzini & Savoia 2011a, 2011f; Manzini et al. 2016). Thus middle voice is associated with several externalizations (a dedicated inflection, a clitic, the *be* auxiliary in Albanian) and with several non-truth equivalent interpretations (passive, reflexive, anticausative, impersonal). Each exponent is compatible with all interpretations and each interpretation can be coupled with any of the exponents. The standard solution to these facts is to assume that the different interpretations are associated with different syntaxes – and that the matter of choosing the right exponent (according to tense and aspect) is demanded to a realizational DM-like externalization component. Savoia and colleagues argue that this is mere encoding of the facts and that a true explanation must embrace the many-to-many mapping between meaning and externalization.

The contributions to the 'Micro- and Macro-variation' part of this book cover many of the themes just raised. Baldi and Franco's contribution has been

discussed in Section 2. Manzini's contribution in turn takes up enclisis/proclisis alternations, as briefly discussed in Section 2, and compares them to pronoun allomorphies governed by irrealis modality that have been prominently discussed in the typological literature (Mauri & Sansò 2012 and references quoted there). The typological debate focusses on whether an irrealis category can be defined faced to the subtle variation observed the diverse languages considered (treatment of negation etc.). Yet that some cluster of modal properties govern the relevant alternations is generally accepted. Manzini makes the point that familiarity breeds low-level solutions to the Romance data, where treatment internal to the morphophonology are routine. The approach of Manzini and Savoia (2017c) however in connecting them directly to the externalization of modal contexts reveals the relevance of microvariation for macrotypologies (see also the conclusions of the contribution by Loporcaro et al.).

Two articles in this volume are devoted to negation. Poletto and Olivieri consider the Occitan variety of L'Escarène and two Veneto varieties (Zemignana and Venice). In Escareac, the preverbal negation *noun* only occurs in negative imperatives, with present and past subjunctive, future indicative, conditional, i.e. irrealis forms. In Venetian, *nɔ* only occurs in (non periphrastic) negative imperatives, in the Zemignana variety strict negative concord (i.e. in the presence of preverbal negations) characterizes modal environments. The [-realis] property therefore unifies the otherwise different lexical alternations in these languages. Cocchi turns to the externalization of the negation in Bantu languages, specifically Swahili and Tchiluba. Negation is expressed with at least three different types of morphemes, a prefix preceding the subject prefix ('pure' negation); Tense/Aspect infixes (or copulas), which follow the subject prefix (aspectual negation); a final inflection (modal negation). She then goes on to compare this distribution with that of negations in Romance microvariation which also can precede or follow the subject clitic or be indeed in sentence final position.

Schifano and Sheehan consider Italian causative constructions. Cinque's (2004) monoclausal analysis predicts that a *volere* modal cannot be embedded under a causative verb, however Schifano and Sheehan show that this is possible when the cause is cliticized to the causative verb. They argue that full dative causees intervene on the agreement path between *fare* and the embedded object but not dative clitics. Turano contributes a work on Albanian in comparison also with other Balkan languages. Albanian displays raising from finite clauses selected by the modal *duket* 'seem', which may agree with the raised DP or keep its non-agreeing (3SG) form. Because of the obvious technical difficulties with raising from finite clauses in standard minimalism, Turano concludes that the movement is in reality a topicalization (with the invariable matrix verb) or else a sideward movement. Still on (non-)finite embedded sentences is the contribution of Cinque and

Benincà on infinitival relatives, of which they point out several properties. These include the distinction between two modal values of '*da* + infinitive' relative constructions (roughly 'could' and 'should') and a difference between *da* infinitival relatives and infinitival relatives introduced by relative pronouns (in that only the latter allow a left periphery). They also show that infinitival relatives can be non-restrictive and argue that infinitival relatives display bounded movement, following Burzio (1986) (though I, for one, don't share the judgements).

The left periphery of the clause is targeted in the article by Rizzi and in that by Belletti. Rizzi discusses an empirical generalization involving topic and focus structures: "while topics may proliferate, at most one left peripheral-focus seems to be allowed per clause". He observes that while two left peripheral foci in the same clause are systematically banned across languages, the co-occurrence of left peripheral foci in different clauses is blocked in a language like Italian but allowed in Gungbe, suggesting a PF mapping parameter for this difference. Belletti focusses on data, showing that developing children speaking a variety of Italian without DOM, nevertheless produce *a*-marked topics (in a given experimental setting); she finds that a language like Balearic Catalan provides a close match to the competence displayed by her developing subjects. Belletti proposes to link *a*-marking to topicality, so that the left-periphery is a core domain for the manifestation of *a*-marking; she also proposes that the *a*-marking of subjects surfacing in various Romance varieties is due to their topic(-like) status. Finally, the object of study of Pescarini and Donzelli are clefts and pseudo-clefts in the variety of Comun Nuovo (Bergamo). They point out that in this variety only DPs are accepted in the focal position (not PPs) including temporal DPs. In interrogatives, the distribution is even more restricted since only temporal clefts are accepted. Pescarini and Donzelli explain the restriction against PP foci by assuming that in the Comun Nuovo variety DP clefts are in fact concealed pseudo clefts, independently restricted to DP foci. Therefore at least in some varieties, (apparent) clefts do not depend on focus movement but rather on copular structures.

4.   Language and its speakers

The title of this section borrows that of a series of books by Baldi and Savoia (2009a, 2009b, cf. also 2017c). Languages, as discussed in the previous sections of this introduction, are formal systems, yielding to internal explanations in terms of primitives and rules of the system; to be more precise, in Chomskyan terms, this characterization holds of the knowledge that speaker/hearers have of a given language (I-language rather than E-language). Yet for all of this, languages are also put to use in various communication settings, with different intents and results.

Savoia's work reaches out to the interfacing of languages with extra-linguistic reality – where the Chomskyan paradigm sheds new light on many classical issues in sociolinguistics, language policy, discourse and communication.

Reference to minority languages, spoken in contact with national language and/or other local varieties, has often been made in previous sections and characterizes several contributions to this volume. By contact, we mean any situation in which two different languages are spoken by the same community, hence in condition of bilingualism. Thus so-called Italian dialects are spoken in conditions of bilingualism with standard Italian. Speakers of Italo-Albanian (Arbëresh) varieties are similarly bilingual with Italian and often with local South Italian dialects. The sociolinguistic picture is further complicated by the fact that Arbëresh/Albanian is recognized as an historical minority language (*lingua di minoranza storica*) under Italian law, leading to the usual issues of standardization vs preservation of local dialects.

The work of Savoia and his coauthors touches on many aspects of the complex question posed by linguistic contact, bilingualism, partial competences in the framework of Chomsky's conception of language as a biological object. These include an active involvement in the implementation of minority languages legislation, with particular reference to Arbëresh (Savoia 2002a, 2002b), and a considerable number of volumes specifically devoted to Italo-Albanian dialects not only in their formal aspects, but also with reference (more or less systematic) to language acquisition, mixed competences and in general relevant psycho- and sociolinguistic domains (Savoia 1984a, 1994a, 2003a, 2012; Savoia et al. 2015).

In his work with Baldi, the continued interest in Arbëresh as a test case in bilingual competence comes to embrace in a more explicit way the cultural identitary connotations of the minority language (Baldi & Savoia 2016a, 2016b, 2017b). Non-linguistic factors, linked up with emotional, socio-cultural and identity components, influence the attitude towards the different languages that the speaker has internalized. The different languages and language varieties that one speaks, together with the place where one was born and where one lives, are factors that play a key role in the processes of distinguishing oneself from and identifying with speakers from other language communities. Immigrants' partial competences provide a related focus in this sustained investigation of the identity-forming value of language (Baldi & Savoia 2006). The use of language in a communicative setting is explored specifically in relation to the rhetorical practices of political discourse (delegitimization, Baldi & Savoia 2017b).

The final set of contributions to the volume, collected in the section on 'Language in context' brings these themes into relief and especially highlight their interplay with mentalistic conceptions of language. Two contributions to the volume are lexicographic studies. Altimari chooses the food lexicon, and in particular the word *lucanica* (Italian, 'Lucanian sausage', hence 'sausage'), as his guide in

his wide-ranging exploration of the material and cultural exchanges between the Italian and the Balkan peninsula, including Albanian, but also Greek and Eastern Romance (Romanian and Aromanian). Trumper contributes the etymologies of Albanian *bërrakë* 'bog', *blertë* 'greenish; verdant', *brī* 'horn', *dritë* 'light', concluding from them that Albanian shares a certain number of isoglosses with Celtic (and both with German). This may locate the origin of Albanian (a *vexata quaestio* in Indo-European linguistics) in North-Central Europe. He favors the notion of an Albanoid group (including pre-Roman Messapic in South Italy) – but avoiding the label of Illyrian, noted "for its conceptual paucity". The work by Desideri connects directly to one of the themes briefly evoked in this section. Desideri reviews the literature on political discourse in the years 1960–1980 including studies on the French revolution, on the discourse of totalitarian leaders (Mussolini, Hitler, Franco), on the political language of the two Germanies. This research, characterized by a lexical-semantic approach, is important to the extent also that it underlies current approaches more oriented to pragmatics and communication.

Finally, the studies by Marotta and Grimaldi remind us of the crucial interaction between formal research (in morphology, in phonology) and recently developed experimental methods, involving both impaired and normal individuals. Marotta dwells on several populations, including Specific Language Impairment and Developmental Dyslexia subjects, whose relatedness is a currently debated topic – as well as normally developing L1 learners. For all of these domains, as well as for language evolution, she provides an up-to-date critical assessment of the literature, seen mostly through the magnifying lens of morphology. Grimaldi outlines state-of-the art research coupling traditional dialectological data with the most modern neurolinguistics techniques, specifically ERP (Evoked Related Potentials) methodologies. Phonological issues are at forefront of the work of Grimaldi and his lab, and therefore prominent in his critical review of the literature – neatly closing the circle where we began our discussion, namely (inescapably perhaps) with phonological elements.

References

Baker, Mark. 2008. The macroparameter in a microparametric world. In *The Limits of Syntactic Variation* [Linguistik Aktuell/Linguistics Today 132] Theresa Biberauer (ed.), 351–373. Amsterdam: John Benjamins. https://doi.org/10.1075/la.132.16bak

Belvin, Robert & den Dikken, Marcel. 1997. *There*, happens, *to, be, have*. *Lingua* 101: 151–183. https://doi.org/10.1016/S0024-3841(96)00049-6

Berwick, Robert & Chomsky, Noam. 2011. The biolinguistic program: The current state of its evolution and development. In *The Biolinguistic Enterprise*, AnnaMaria Di Sciullo & Cedric Boeckx (eds), 19–41. Oxford: OUP.

Biberauer, Theresa, Holmberg, Anders & Roberts, Ian. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45: 169–225. https://doi.org/10.1162/LING_a_00153

Blevins, James, P. 2006. Word-based morphology. *Journal of Linguistics* 42: 531–573. https://doi.org/10.1017/S0022226706004191

Borer, Hagit. 2005. *Structuring Sense*, Vol. 1: *In name only*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199263905.001.0001

Burzio, Luigi. 1986. *Italian Syntax*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-4522-7

Calabrese, Andrea. 1998. Some remarks on the Latin case system and its development in Romance. In *Theoretical Advances on Romance Languages* [Current Issues in Linguistic Theory 157] José Lema & Esthela Trevino (eds), 71–126. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.157.06cal

Calabrese, Andrea. 2008. On absolute and contextual syncretism. In *The Bases of Inflectional Identity*, Andrew Nevins & Asaf Bachrach (eds), 156–205. Oxford: OUP.

Chomsky, Noam & Halle, Morris. 1968. *The Sound Pattern of English*. New York NY: Harper & Row.

Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, Stephen Anderson & Paul Kiparsky (eds), 232–286. New York NY: Holt, Reinhart and Winston.

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

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

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, Michael Kenstowicz (ed.), 1–54. Cambridge MA: The MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1–22. https://doi.org/10.1162/0024389052993655

Chomsky, Noam & Berwick, Robert. 2015. *Why Only Us: Language and Evolution*. Cambridge MA: The MIT Press.

Chomsky, Noam, Gallego, Angel & Ott, Dennis. 2018. Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*.

Cinque, Guglielmo. 2004. Restructuring and functional structure. In *Structures and Beyond. The Cartography of Syntactic Structures*, Vol. 3, Adriana Belletti (ed.), 132–191. Oxford: OUP.

Cinque, Guglielmo & Rizzi, Luigi. 2010. The cartography of syntactic structures. In *The Oxford handbook of linguistic analysis*, Bernd Heine & Heiko Narrog (eds), 51–65. Oxford: OUP.

Culicover, Peter & Jackendoff, Ray. 2005. *Simpler Syntax*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199271092.001.0001

D'Alessandro, Roberta & Scheer, Tobias. 2013. Phase head marking. In *Universal Syntax and Parametric Phonology* Hisao Tokizaki & Yoshi Dobashi (eds). *Linguistic Analysis* 38(4): 305–330.

D'Alessandro, Roberta & van Oostendorp, Marc. 2016a. Abruzzese metaphony and the A eater In *Approaches to Metaphony in the Languages of Italy*, Francesc Torres-Tamarit, Kathrin Linke & Marc van Oostendorp (eds), 349–368. Berlin: Mouton de Gruyter.

D'Alessandro, Roberta & van Oostendorp, Marc. 2016b. When imperfections are perfect – Prosody, phi-features and deixis in central and southern Italian vocatives. In *Romance Languages and Linguistic Theory – Selected papers from 'Going Romance' 28, Lisbon*, [Romance Languages and Linguistic Theory 10], Ernestina Carrilho, Alexandra Fiéis, Maria Lobo & Sandra Pereira (eds), 61–82. Amsterdam: John Benjamins. https://doi.org/10.1075/rllt.10.04dal

Déchaine, Rose-Marie, Girard, Raphaël, Mudzingwa, Calisto & Wiltschko, Martina. 2014. The internal syntax of Shona class prefixes. *Language Sciences* 43: 18–46. https://doi.org/10.1016/j.langsci.2013.10.008

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, Kenneth Hale & Samuel J. Keyser (eds), 111–176. Cambridge MA: The MIT Press.

Halle, Morris & Vaux, Bert. 1998. Theoretical aspects of Indo-European nominal morphology: The nominal declensions of Latin and Armenian. In *Mír Curad: Studies in Honor of Calvert Watkins*, Jay Jasanoff, H. Craig Melchert & Lisi Oliver (eds), 223–240. Innsbruck: Innsbrucker Beitraege zur Sprachwissenschaft.

Hauser, Mark D., Chomsky, Noam & Fitch, W. Tecumseh. 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298: 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16: 547–621.

Kaye, Jonathan, Lowenstamm, Jean & Vergnaud, Jean-Roger. 1985. The internal structure of phonological elements: A theory of charm and government, *Phonology Yearbook* 2: 302–329. https://doi.org/10.1017/S0952675700000476

Kaye, Jonathan, Lowenstamm, Jean & Vergnaud, Jean-Roger. 1990. Constituent Structure and Government in Phonology, *Phonology* 7: 193–231. https://doi.org/10.1017/S0952675700001184

Kayne, Richard, S. 2010. *Comparisons and Contrasts*. Oxford: OUP.

Mauri, Caterina & Sansò, Andrea. 2012. What do languages encode when they encode reality status? *Language Sciences* 34: 99–106. https://doi.org/10.1016/j.langsci.2010.11.004

Newell, Heather & Piggott, Glyne. 2014. Interactions at the syntax-phonology interface: Evidence from Ojibwe, *Lingua* 150: 332–362. https://doi.org/10.1016/j.lingua.2014.07.020

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge MA: The MIT Press https://doi.org/10.7551/mitpress/9780262014304.001.0001

Leonardo M. Savoia complete bibliography 1974–2017

Baldi, Benedetta & Savoia, Leonardo M. 2017a. Le vocali toniche nelle varietà friulane, ladine e romance. *Ladinia* 53–80.

Baldi, Benedetta & Savoia, Leonardo M. 2017b. L'identità come mezzo di (de)legittimazione: Procedure linguistiche e pragmatiche nel discorso politico. In *La delegittimazione politica nell'età contemporanea, 2: Parole nemiche: teorie, pratiche e linguaggi*, Benedetta Baldi (ed.), 287–338. Roma: Viella.

Baldi, Benedetta & Savoia, Leonardo M. 2017c. *Linguaggio & comunicazione. Introduzione alla linguistica*. Bologna: Zanichelli.

Baldi, Benedetta & Savoia, Leonardo M. 2017d. Cultura e identità nella lingua albanese. *LEA* 6: 45–66.

Baldi, Benedetta & Savoia, Leonardo M. 2017e. Il significato in Saussure e nella linguistica europea del novecento. In *Ce qui nous est donné, ce sont les langues Studi linguistici in onore di Maria Pia Marchese*, Monica Ballerini, Francesca Murano & Letizia Vezzosi (eds) 197–213. Alessandria: Edizioni dell'Orso.

Manzini, Maria Rita & Savoia, Leonardo M. 2017a. N morphology and its interpretation: The neuter in Italian and Albanian varieties. In *Constraints on Structure and Derivation in Syntax, Phonology and Morphology*, Anna Bloch-Rozmej & Anna Bondaruk (eds), 213–236. Frankfurt: Peter Lang.

Manzini, Maria Rita & Savoia, Leonardo M. 2017b. Gender, number and inflectional class in Romance: Feminine/plural -*a*. In *Language Use and Linguistic Structure*, Joseph Emonds & Markéta Janebova (eds), 263–282. Olomouc: Palacky University Olomouc.

Manzini, Maria Rita & Savoia, Leonardo M. 2017c. Enclisis/proclisis alternations in Romance: Allomorphies and (re)ordering. *Transactions of the Philological Society* 115: 98–136. https://doi.org/10.1111/1467-968X.12093

Manzini, Maria Rita & Savoia, Leonardo M. 2017d. Case in the nominal and pronominal systems in Aromanian: Oblique case and its interactions with the person split. *Revue Roumaine de Linguistique* LXII: 115–142.

Manzini, Maria Rita, Lorusso, Paolo & Savoia, Leonardo M. 2017. a/bare finite complements in Southern Italian varieties: Mono-clausal or bi-clausal syntax? *QULSO* 3: 11–59.

Savoia, Leonardo M. & Baldi, Benedetta. 2017. Enhancing stressed /a/ low frequency components in the context of sonorants. In *Constraints on Structure and Derivation in Syntax, Phonology and Morphology*, Anna Bloch-Rozmej & Anna Bondaruk (eds), 237–263. Frankfurt: Peter Lang.

Savoia, Leonardo M., Manzini, M. Rita, Baldi, Benedetta & Franco, Ludovico. 2017a. A morphosyntactic analysis of evaluatives in Italian. *SILTA* XLVI: 413–440.

Savoia, Leonardo M., Manzini, M. Rita, Franco, Ludovico & Baldi, Benedetta. 2017b. Nominal evaluative suffixes in Italian. In *Di tutti i colori. Studi linguistici per Maria Grossmann*, Roberta d'Alessandro, Gabriele Iannàccaro, Diana Passino & Anna M. Thornton (eds), 283–300. Utrecht: Utrecht University Repository.

Baldi, Benedetta & Savoia, Leonardo M. 2016a. Le differenze linguistiche e le lingue minoritarie nella società attuale. *SILTA* XLV: 211–243.

Baldi, Benedetta & Savoia, Leonardo M. 2016b. Fenomeni di code-mixing e di prestito nei sistemi arbëreshë. In *La lingua e i parlanti. Studi e ricerche di linguistica*, Benedetta Baldi & Leonardo M. Savoia (eds). 145–163. Alessandria: Edizioni dell'Orso.

Baldi, Benedetta & Savoia, Leonardo M. 2016c. Il formalismo russo, Roman Jakobson e la linguistica nella prima metà del Novecento. *LEA* 5: 585–621.

Franco, Ludovico, Manzini M. Rita & Savoia, Leonardo M. 2016. N morphology and its interpretation: The neuter in Central Italian varieties. In *La lingua e i parlanti. Studi e ricerche di linguistica*, Benedetta Baldi & Leonardo M. Savoia (eds), 51–78. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo M. 2016a. Which syntactic information gets interpreted by phonosyntactic processes? Evidence from /u/ propagation in Southern Italian dialects. In *Approaches to Metaphony in the Languages of Italy*, Francesc Torres-Tamarit, Kathrin Linke & Marc vanOostendorp (eds), 222–258. Berlin: De Gruyter.

Manzini, M. Rita & Savoia, Leonardo M. 2016b. Caso e proprietà denotazionali: La flessione nominale nella varietà ghega di Shkodër. In *La lingua e i parlanti. Studi e ricerche di linguistica*, Benedetta Baldi & Leonardo M. Savoia (eds), 79–93. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita, Anna, Roussou & Savoia, Leonardo M. 2016. Middle-passive voice in Albanian and Greek. *Journal of Linguistics* 52: 111–150. https://doi.org/10.1017/S0022226715000080

Savoia, Leonardo M., & Benedetta, Baldi. 2016a. Un contributo alla ricostruzione del vocalismo tonico toscano: Le vocali toniche medio-basse. In *La lingua e i parlanti. Studi e*

ricerche di linguistica, Benedetta Baldi & Leonardo M. Savoia (eds), 223–242. Alessandria: Edizioni dell'Orso.

Savoia Leonardo, M. & Baldi, Benedetta. 2016b. Proprietà morfo-sintattiche dei suffissi valutativi in italiano. In *La lingua e i parlanti. Studi e ricerche di linguistica*, Benedetta Baldi & Leonardo M. Savoia (eds), 95–113. Alessandria: Edizioni dell'Orso.

Savoia, Leonardo M. & Baldi, Benedetta. 2016c. Armonie vocaliche e metafonia nelle varietà Siciliane. *Bollettino – Centro di Studi Filologici e Linguistici Siciliani* 27: 201–237.

Savoia, Leonardo M. & Baldi, Benedetta. 2016d. Propagation and preservation of rounded back vowels in Lucanian and Apulian varieties. *QULSO* 2: 11–58.

Savoia, Leonardo M. & M. Rita Manzini. 2016. Quale informazione sintattica viene interpretata dai fenomeni fonosintattici? Evidenza dalle varietà italiane meridionali. In *Congrès international de linguistique et de philologie romanes, Nancy, 15–20 luglio 2013*, 403–414. Nancy: ATILF.

Savoia, Leonardo M. 2016. Harmonic processes and metaphony in some Italian varieties. In *Approaches to Metaphony in the Languages of Italy*, Francesc Torres-Tamarit, Kathrin Linke, & Marc van Oostendorp (eds), 9–53. Berlin: De Gruyter.

Franco, Ludovico, Manzini, M. Rita & Savoia, Leonardo M. 2015a. Linkers and agreement. *The Linguistic Review* 32: 277–332.

Franco Ludovico, Manzini M. Rita, & Savoia, Leonardo M. 2015b. N morphology and its interpretation: The neuter in Central Italian varieties and its implications. *Isogloss* 2: 41–67. https://doi.org/10.5565/rev/isogloss.14

Manzini, M. Rita, Savoia, Leonardo M. & Franco, Ludovico. 2015. Ergative case, aspect and person splits: Two case studies. *Acta Linguistica Hungarica* 62: 297–351. https://doi.org/10.1556/064.2015.62.3.3

Manzini, M. Rita & Savoia, Leonardo M. 2015a. Notes on the contact between Italo-Albanian and Romance (Calabrian, Lucanian) varieties: Borrowings, code-mixing and convergence. *Hylli I Dritës* 35: 92–116.

Manzini, M. Rita & Savoia, Leonardo M. 2015b. Oblique case and Indo-European ergativity splits. *Revue Roumaine de Linguistique*. LX: 297–315.

Manzini, M. Rita & Savoia, Leonardo M. 2015c. Causatives and 'inverse' agreement in Ginestra (Arbëresh). In *Studime Në nderim të prof. Francesco Altimarit me rastin e 60-vjetorit të lindjes Studi In onore del prof. Francesco Altimari in occasione del 60° compleanno*, Bardhil Demiraj, Matteo Mandalà & Shaban Sinani (eds), 345–364. Tirana: Albpaper.

Savoia, Leonardo M. 2015a. *I dialetti italiani. Sistemi e processi fonologici nelle varietà di area italiana e romancia*. Pisa: Pacini.

Savoia, Leonardo M. 2015b. La minoranza linguistica arbëreshe. In *Migrazioni tra disagio linguistico e patrimoni culturali/ Les migrations entre malaise linguistique et patrimoines culturels – Actes des Sixièmes Journées des Droits Linguistiques*, Giovanni Agresti & Silvia Pallini (eds), 253–278. Roma: Aracne.

Savoia, Leonardo M. 2015c. La linguistica generativa. In *Scienze del linguaggio e educazione linguistica*, Michele Daloiso (ed.), 21–32, Torino: Loescher.

Savoia, Leonardo M., Manzini, M. Rita & Baldi, Benedetta. 2015. *Studime në gjuhësinë shqiptare*. Prishtinë: Akademia e Shkencave dhe Arteve e Kosovës – Seksioni i Gjuhësisë dhe i Letërsisë.

Savoia, Leonardo M. & Manzini, M. Rita. 2015. L'ordre des clitiques objets dans les variétés corses. In *Paroddi varghji. Mélanges offerts à Marie-José Dalbera-Stefanaggi*, Stella Retali-Medori (ed.), 369–394. Alessandria: Edizioni dell'Orso.

Savoia, Leonardo M. & Vinciguerra, Antonio. 2015. Appunti di storia della linguistica italiana: Il contributo fiorentino. *LEA* 4: 41–78.

Manzini, M. Rita & Savoia, Leonardo M. 2014a. From Latin to Romance: case loss and preservation in pronominal systems. *Probus* 26: 217–248.

Manzini, M. Rita & Savoia, Leonardo M. 2014b. Linkers in Aromanian in comparison to Albanian and Romanian. *Rivista di Grammatica Generativa* 36: 83–104,

Manzini, M. Rita & Savoia, Leonardo M. 2014c. Person splits in the case systems of Geg Albanian (Shkodër) and Arbëresh (Greci). *SILTA* XLIII: 7–42.

Manzini, M. Rita, Franco, Ludovico & Savoia, Leonardo M. 2014. Linkers Are Not 'Possession Markers' but 'Agreements'. In *Complex Visibles out there*, Ludmila Veselovská & Markéta Janebová (eds), 233–251, Olomouc: Palacký University Press.

Savoia, Leonardo Maria. 2014. *Introduzione alla fonetica e alla fonologia*. Bologna: Zanichelli.

Manzini, M. Rita & Savoia, Leonardo M. 2012a. 'Case' categories in the Geg Albanian variety of Shkodër. *Res Albanicae* 1: 23–42.

Manzini, M. Rita & Savoia, Leonardo M. 2012b. On the lexical/ functional divide: The case of negation. In *Functional Heads*, Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds), 338–350. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199746736.003.0026

Manzini, M. Rita & Savoia, Leonardo M. 2012c. Non-active voice in Albanian: Implications for the theory of movement. In *Studi sulle varietà arbëreshe*, Leonardo M. Savoia (ed.), 111–149. Rende: Università della Calabria.

Manzini, M. Rita & Savoia, Leonardo M. 2012d. Sentential negation in Piedmontese varieties. In *Studi sui dialetti del Piemonte*, Jacopo Garzonio (ed.), 123–145. Padua: Unipress.

Manzini, M. Rita & Savoia, Leonardo M. 2012e. The interaction of DOM and discourse-linking in Arbëresh pronouns: Case as nominal property. In *Theoretical and empirical studies in Albanian Syntax. Studime teorike dhe empirike në fushën e sintaksës së shqipes*, Giuseppina Turano, Rami Memushaj & Flora Koleci (eds), 119–139, *Munich: Lincom.*

Savoia, Leonardo M. & Manzini, M. Rita. 2012a. Morphosyntax of the participle and infinitive in the variety of Shkodër. In *Studime për nder të Rexhep Ismajlit me rasin e 65-vjetorit të lindjes*, Bardh Rugova (ed.), 651–677. Prishtinë: Koha.

Savoia, Leonardo M. & Manzini, M. Rita. 2012b. Case as nominal property: The interaction of DOM and discourse-linking in Shkodër pronouns. In *Shqipja dhe gjuhët e Ballkanit – Albanian and Balkan Languages*, Rexhep Ismajli (ed), 411–432. Prishtinë: Akademia e Shkencave dhe e Arteve e Kosovës – Akademia e Shkencave e Shqipërisë.

Savoia, Leonardo M. 2012. *Studi sulle varietà arbëreshe*, 2nd ed. Rende: Università della Calabria, Dipartimento di Linguistica, Sezione di Albanologia.

Manzini, M. Rita & Savoia, Leonardo M. 2011a. *Grammatical Categories: Variation in Romance Languages*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511974489

Manzini, M. Rita & Savoia, Leonardo M. 2011b. Reducing 'case' to denotational primitives: Nominal inflections in Albanian. *Linguistic Variation* 11: 76–120. https://doi.org/10.1075/lv.11.1.03man

Manzini, M. Rita & Savoia, Leonardo M. 2011c. Mesoclisis in the imperative: Phonology, morphology or syntax? *Lingua* 121: 1101–1120. https://doi.org/10.1016/j.lingua.2011.02.002

Manzini, M. Rita & Savoia, Leonardo M. 2011d. (Definite) denotation and case in Romance. History and variation. In *Romance Languages and Linguistic Theory 2009. Selected Papers from 'Going Romance' Nice 2009* [Romance Languages and Linguistic Theory 3] Janine

Berns, Haike Jacobs & Tobias Scheer (eds), 149–165. Amsterdam: John Benjamins. https://doi.org/10.1075/rllt.3.09man

Manzini, M. Rita & Savoia, Leonardo M. 2011e. Wh-in situ & wh-doubling in Northern Italian varieties: Against remnant movement. *Linguistic Analysis* 37: 79–113.

Manzini, M. Rita & Savoia, Leonardo M. 2011f. (Bio)linguistic variation: Have/ be alternations in the present perfect. In *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, Anna Maria Di Sciullo & Cedric Boeckx (eds), 222–265. Oxford: OUP.

Savoia, Leonardo M. 2011. Guerra fredda. In *Itabolario. L'Italia unita in 150 parole*, Massimo Arcangeli (ed.), 213–215. Roma: Carocci.

Baldi, Benedetta & Savoia, Leonardo M. 2010. Metafora e ideologia nel linguaggio politico. In *Lingua Italiana d'Oggi*, Massimo Arcangeli (ed), 119–165. Roma: Bulzoni.

Manzini, M. Rita & Savoia, Leonardo M. 2010a. Case as denotation: Variation in Romance. *SILTA* XXXIX: 409–438.

Manzini, M. Rita & Savoia, Leonardo M. 2010b. Syncretism and suppletion in clitic systems: Underspecification, silent clitics or neither? In *Syntactic Variation. The Dialects of Italy*, Roberta D'Alessandro, Adam Ledgeway & Ian Roberts (eds), 86–101, Cambridge: CUP.

Manzini, M. Rita & Savoia, Leonardo M. 2010c. Rasa dhe veti treguese: Emëror në të folmen gege të Shkodrës. In *Identitete gjuhësore dhe kulturore*, 13–29. Shkodër: Universiteti i Shkodrës 'Luigj Gurakuqi'.

Savoia, Leonardo M. & Manzini, M. Rita. 2010a. Lexicalization of 3rd person object clitics: Clitic enclisis and clitic drop. In *Syntactic Variation. The Dialects of Italy*, Roberta D'Alessandro, Adam Ledgeway & Ian Roberts (eds), 102–118. Cambridge: CUP.

Savoia, Leonardo M. & Manzini, M. Rita. 2010b. Sintagma nominale e caso a Shkodër. *Hylli I Dritës* 3: 111–122.

Savoia, Leonardo M. & Manzini, M. Rita. 2010c. Les clitiques sujets dans les variétés occitanes et francoprovençales italiennes. *Corpus* 9: 165–189.

Savoia, Leonardo M. 2010a. Albanese, comunità. In *Enciclopedia dell'italiano*, 2 Vols, Raffaele Simone, Gaetano Berruto & Paolo D'Achille (eds), 33–37. Roma: Istituto dell'Enciclopedia Italiana G. Treccani.

Savoia, Leonardo M. 2010b. Prestiti lessicali e code-mixing nei sistemi arbëreshë. In *Parole. Il lessico come strumento per organizzare e trasmettere gli etnosaperi*, Nadia Prantera, Antonio Mendicino & Cinzia Citraro (eds), 717–738. Rende: Università della Calabria.

Savoia, Leonardo M. 2010c. Diritti linguistici. In *Enciclopedia dell'italiano*, 2 Vols, Raffaele Simone, Gaetano Berruto & Paolo D'Achille (eds), 374–376. Roma: Istituto dell'Enciclopedia Italiana G. Treccani.

Baldi, Benedetta & Savoia, Leonardo M. 2009a. *Lingua e comunicazione. La lingua e i parlanti*. Pisa: Pacini.

Baldi, Benedetta & Savoia, Leonardo M. 2009b. *Lingua e società. La lingua e i parlanti*. Pisa: Pacini.

Baldi, Benedetta & Savoia, Leonardo M. 2009c. Semantica delle lingue naturali: una breve rassegna. *Quaderni del Dipartimento di Linguistica. Università degli Studi di Firenze* 19: 115–142.

Manzini, M. Rita & Savoia, Leonardo M. 2009a. Mesoclisis in the Imperative: Phonology, Morphology or Syntax? In *Proceedings XXXV Incontro di Grammatica Generativa*, Vincenzo Moscati & Emilio Servidio (eds), 51–76. Siena: CISCL Centro interdipartimentale di studi cognitivi sul linguaggio.

Manzini, M. Rita & Savoia, Leonardo M. 2009b. Morphology dissolves into syntax: Infixation and doubling in Romance languages. *Annali Online dell'Università di Ferrara. Sezione lettere* IV: 1–28.

Savoia, Leonardo M. 2009. Variazione linguistica e bilinguismo: La mescolanza linguistica nell'arbëresh di Ginestra. In *Alloglossie e comunità alloglotte nell'Italia contemporanea. Teorie, applicazioni e descrizioni, prospettive*, Carlo Consani et al. (eds), 121–141. Roma: Bulzoni.

Baldi, Benedetta & Savoia, Leonardo M. 2008. Identità e variazione linguistica nel quadro dei fenomeni di globalizzazione. In *Lingue e culture fra identità e potere*, Massimo Arcangeli & Carla Marcato (eds), 133–144. Roma: Bonacci.

Manzini, M. Rita & Savoia, Leonardo M. 2008a. Work notes on Romance morphosyntax. Appunti di morfosintassi romanza. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo M. 2008b. Uninterpretable features are incompatible in morphology with other minimalist postulates. In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, Robert Freidin, Carlos Otero & Maria Luisa Zubizarreta (eds), 43–72. Cambridge MA: The MIT Press.
https://doi.org/10.7551/mitpress/9780262062787.003.0004

Savoia, Leonardo M. & Carpitelli, Elisabetta. 2008. Problèmes de micro-variation phonologique dans les domaines dialectaux de l'Italie septentrionale. *Revue Française de Linguistique Appliquée* 13: 103–119.

Savoia, Leonardo M. 2008a. *Studi sulle varietà arbëreshe*. Cosenza: Università della Calabria.

Savoia, Leonardo M. 2008b. La linguistica di Paolo Marzolo e il pensiero scientifico del suo tempo. *SILTA* 37: 511–548.

Savoia, Leonardo M. 2008c. Marzolo, Paolo. *Istituto della Enciclopedia Italiana. Dizionario Biografico degli Italiani* LXXI: 486–488.

Baldi, Benedetta & Savoia, Leonardo M. 2007a. Linguaggio e globalizzazione. In *Istituto della Enciclopedia Italiana. Enciclopedia Italiana – XXI secolo – VII Appendice*, 286–287, Roma: Istituto della Enciclopedia Italiana.

Baldi, Benedetta & Savoia, Leonardo M. 2007b. Prospettive di educazione linguistica. In *GISCEL. Educazione linguistica democratica. A trent'anni dalle Dieci Tesi*, 168–196. Milano: Franco Angeli.

Manzini, M. Rita & Savoia, Leonardo M. 2007. *A Unification of Morphology and Syntax. Investigations into Romance and Albanian dialects*. London: Routledge.

Savoia, Leonardo M. & M. Rita Manzini. 2007. Variazione sintattica nel costrutto ausiliare arbëresh. La variazione come problema teorico. In *Minoranze linguistiche. Prospettive, strumenti, territori*, Carlo Consani & Paola Desideri (eds), 85–102. Roma: Carocci.

Savoia, Leonardo M. & Manzini, M. Rita. 2007. Fenomeni sintattici romanci. In *Ladine Loqui – IV Colloquium Retoromanistich*, Federico Vicario (eds), 13–63. Udine: Società Filologica Friulana.

Savoia, Leonardo M. 2007a. Clitici oggetto silenti in varietà romanze. *SILTA* 35: 427–469.

Savoia, Leonardo M. 2007b. Introduzione. Aspetti della linguistica di De Rada nel quadro delle ricerche linguistiche arbëreshe del '700 e dell'800. In *Gerolamo De Rada, Opera omnia, X – Opere grammaticali*, Franco Altimari (ed.), 9–26. Soveria Mannelli: Rubbettino.

Savoia, Leonardo M. 2007c. Alcune caratteristiche morfosintattiche della lingua piemontese. In *25 ani 'd piemuntèis a Muncalv*, Giuseppe Prosio & Piero Baldovino (eds), 47–84. Asti: Casa Editrice Espansione Grafica.

Savoia, Leonardo M. 2007d. Recensione a Laura Bafile 'Le consonanti finali nel fiorentino e nel napoletano'. *Rivista Italiana di Dialettologia* 31: 374–375.

Savoia, Leonardo M. 2007e. La dialettologia italiana: Il contributo fiorentino. In *Firenze e la lingua italiana fra Nazione e Europa*. Nicoletta Maraschio (ed.), 125–154. Firenze: Firenze University Press.

Baldi, Benedetta & Savoia, Leonardo M. 2006. *Perché barbari? Lingua, comunicazione e identità nella società globale*. Roma: Bulzoni.

Savoia, Leonardo M. 2006a. Elementi di linguistica albanese. In *Albanesi, cossovari e molti altri nella scuola di Monteroni D'Arbia*, Iesa (SI), Cinzia Fia (ed.), 75–100. Edizioni Gorée.

Savoia, Leonardo M. 2006b. Il lavoro linguistico di Gastone Venturelli. *Lares* LXX: 373–385.

Baldi, Benedetta & Savoia, Leonardo M. 2005a. I media e la formazione dell'opinione pubblica. Alcune riflessioni sul rapporto tra informazione e globalizzazione. *Quaderni del dipartimento di linguistica. Università degli studi di Firenze* 15: 255–279.

Baldi, Benedetta & Savoia, Leonardo M. 2005b. Mezzi di comunicazione di massa e scuola: Conflitto d'interessi? media, informazione e educazione linguistica in Italia. In *Lingua Italiana D'oggi 2*, Massimo Arcangeli (ed.), 215–268. Roma: Bulzoni.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni Dell'Orso.

Savoia, Leonardo M. 2005. Armonizzazioni vocaliche in alcune varietà romanze. In *Italia linguistica: discorsi di scritto e di parlato. Nuovi studi di linguistica italiana in onore di Giovanni Nencioni*, Marco Biffi, Omar Calabrese & Luciana Salibra (eds), 217–233. Siena: Protagon.

Manzini, M. Rita & Savoia, Leonardo M. 2004a. Clitics: Cooccurrence and mutual exclusion patterns. In *The Structure of CP and IP*, Luigi Rizzi (ed.), 211–250. Oxford: OUP.

Manzini, M. Rita & Savoia, Leonardo M. 2004b. The nature of the agreement inflections of the verb. In *Romance op. 47. Collected Papers on Romance Syntax* [MIT Working Papers in Linguistics 47], Ana Castro et al. (eds), 149–178. Cambridge MA: MITWPL.

Manzini, M. Rita & Savoia, Leonardo M. 2004c. Uninterpretable features are incompatible in morphology with other minimalist postulates. *Quaderni del Dipartimento di Linguistica. Università degli Studi di Firenze* 14: 13–33.

Savoia, Leonardo M. 2004a. Presentazione alla seconda edizione. In *La Sapienze De Le Revetole (Sapienza Rubastina). Proverbi In Lingua Vernacola*, Nicola Stragapede (ed.), 13–17. Ruvo di Puglia: Comune di Ruvo.

Savoia, Leonardo M. 2004b. La lingua si difende da sé. In *Lingua Italiana D'oggi 1*, Massimo Arcangeli (ed.), 31–53. Roma: Bulzoni.

Savoia, Leonardo M. 2004c. Etimologji dhe gjuhë kombëtare. Shënime mbi një aspekt të gjusësisë së Eqrem Çabejt. In *Eqrem Çabej dhe kultura shqiptare*, Thoma Dhima & Vasil Bici (eds), 79–97. Tirana: Universiteti 'Eqrem Çabej' Gjirokastër.

Manzini, M. Rita & Savoia, Leonardo M. 2003a. The nature of complementizer. *Rivista Di Grammatica Generativa* 28: 87–110.

Manzini, M. Rita & Savoia, Leonardo M. 2003b. Pjesorja dhe paskajorja në të folmen e Shkodrës. *Studime Filologjike* 2003: 93–138.

Manzini, M. Rita & Savoia, Leonardo M. 2003. Participio e infinito nella varietà di Scutari. In *XXVIII Congresso Internazionale di Studi Albanesi, Università Di Palermo*, 16–19 maggio 2002, M. Mandalà & A. C. Mirror (eds), 401–432.

Savoia, Leonardo M. 2003a. Ideologia nazionale e indagine linguistica. Due conferenze albanesi. *Quaderni del Dipartimento di Linguistica. Università degli Studi di Firenze* 13: 127–152.

Savoia, Leonardo M. 2003b. Introduzione. In *Nuove ricerche sul vocalismo tonico del Salento Meridionale*, Mirko Grimaldi (ed.), xii–xvi. Alessandria: Edizioni Dell'Orso.

Savoia, Leonardo M. 2003c. *I dialetti italo-albanesi e l'Albania*. Tirana: Istituto Italiano di Cultura di Tirana.

Manzini, M. Rita & Savoia, Leonardo M. 2002a. Clitics: Lexicalization patterns of the so-called 3rd person dative. *Catalan Journal of Linguistics* 1: 117–155. https://doi.org/10.5565/rev/catjl.57

Manzini, M. Rita & Savoia, Leonardo M. 2002b. Parameters of subject inflection in Italian dialects. In *Subjects, Expletives and the EPP*, Peter Svenonius (ed.), 157–199. Oxford: OUP.

Manzini, M. Rita & Savoia, Leonardo M. 2002c. Negative adverbs are neither ADN nor NEG. In *Proceedings of the Northeast Linguistic Society (NELS) 32*, Masako Hirotani et al. (eds), 327–346. Amherst MA: GLSA.

Savoia, Leonardo M. 2002a. Componenti ideologiche nel dibattito sulle leggi di tutela linguistica. *Plurilinguismo* 9: 85–114.

Savoia, Leonardo M. 2002b. La legge 482 sulle minoranze linguistiche storiche. Le lingue di minoranza e le varietà non standard in Italia. *Rivista Italiana di Dialettologia* 25: 7–50.

Fici, Francesca, M. Rita Manzini & Savoia, Leonardo M. 2001. Clitics in Macedonian. In *Current Issues in Formal Slavic Linguistics*, Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds), 148–158. Frankfurt: Peter Lang.

Manzini, M. Rita & Savoia, Leonardo M. 2001. The syntax of object clitics: *Si* in Italian dialects. In *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*, Guglielmo Cinque & Giampaolo Salvi (eds), 233–264. Amsterdam: North Holland.

Savoia, Leonardo M. 2001. Edukimi gjuhësor në gjuhën e pakicës. In *Kerkime universitare 5*. AA.VV, 85–109. Gjirokaster: Universiteti i Gjirokastres.

Manzini, M. Rita & Savoia, Leonardo M. 2000. Variazione linguistica, disturbi del linguaggio e grammatica universale. *Quaderni del Dipartimento di Linguistica. Università degli Studi di Firenze* 10: 13–30.

Savoia, Leonardo M. 2000a. Note sulla formazione degli studi linguistici e dialettologici in Italia. *Studi di Grammatica Italiana* 19: 363–421.

Savoia, Leonardo M. 2000b. I dialetti, le lingue di minoranza e la lingua standard nella cultura italiana. In *Le comunità italo-albanesi fra microstoria e arbëreshe: Il caso di San Marzano*, V. A. Martella & G. Carducci (eds), 17–48. S. Marzano: Scuola Media di S. Marzano.

Savoia, Leonardo M. 2000c. L'educazione linguistica in una lingua minoritaria: l'Arbëresh di Piana. In *Skanderbeg 3000. Atti del seminario*, P. Manali (ed.). Piana degli Albanesi: Comune di Piana degli Albanesi.

Savoia, Leonardo M., Manzini, M. Rita & Mancini, Alberto. 2000. Linguistica. In *Enciclopedia Italiana di scienze, lettere ed arti – Appendice 2000*, AA.VV., 85–90. Roma: Enciclopedia Italiana Treccani.

Fici, Francesca, Manzini, M. Rita & Savoia, Leonardo M. 1999. Clitics in Macedonian. *Quaderni Del Dipartimento di Linguistica. Università degli Studi di Firenze* 9: 13–30.

Manzini, M. Rita & Savoia, Leonardo M. 1999. The syntax of middle-reflexive and object clitics: A case of parametrization in Arberesh dialects. In *Studi in onore di Luigi Marlekaj*, M. Mandalà (ed.), 283–328. Bari: Adriatica.

Mancini, Alberto, Manzini, M. Rita & Savoia, Leonardo M. 1998. Linguistica. In *Enciclopedia Italiana – Appendice 2000*, 85–90. Roma: Istituto della Enciclopedia Italiana.

Manzini, M. Rita & Savoia, Leonardo M. 1998a. Clitics and auxiliary choice in Italian dialects: Their relevance for the person ergativity split. *Recherches Linguistiques de Vincennes* 27: 115–138.

Manzini, M. Rita & Savoia, Leonardo M. 1998b. Varietà romanze a 'Verbo Secondo'. In *Echi di memoria. Scritti di varia filologia, critica e linguistica*, Gaetano Chiappini (ed.), 477–509. Firenze: Alinea.

Manzini, M. Rita & Savoia, Leonardo M. 1998c. Negation parameters and their interaction in Italian dialects. *Quaderni di lavoro ASIS* 2: 39–60.

Manzini, M. Rita & Savoia, Leonardo M. 1997. Null subjects without pro. *UCL Working Papers in Linguistics* 11: 303–313

Savoia, Leonardo M. 1997a. Il vocalismo a tre gradi dell'area calabro-lucana. In *Studi linguistici offerti a Gabriella Giacomelli dagli amici e dagli allievi*, AA.VV. (eds), 363–375. Padova: Unipress.

Savoia, Leonardo M. 1997b. Inflectional morphology of the verb. In *The Dialects of Italy*, Martin Maiden & Mair Parry (eds), 75–86. London: Routledge.

Savoia, Leonardo M. 1997c. The geographical distribution of the dialects. In *The Dialects of Italy*, Martin Maiden & Mair Parry (eds), 225–234. London: Routledge.

Savoia, Leonardo M. & Maiden, Martin. 1997. Metaphony. In *The Dialects of Italy*, Martin Maiden & Mair Parry (eds), 15–25. London: Routledge.

Cocchi, Gloria, Giusti, Mariangela, Manzini, M. Rita, Mori, Tiziana & Savoia, Leonardo M. 1996. *L'italiano come L2 nella scuola dell'obbligo*. Roma: Bulzoni,

Savoia, Leonardo M. 1996a. Ascoli. In *Dizionario di storiografia*, A. De Bernardi & S. Guarracino (eds), 83–84. Milano: Bruno Mondadori.

Savoia, Leonardo M. 1996b. Dialettologia. In *Dizionario di storiografia*, A. De Bernardi & S. Guarracino (eds), 305–306. Milano: Bruno Mondadori.

Savoia, Leonardo M. 1996c. Sociolinguistica. In *Dizionario di storiografia*, A. De Bernardi & S. Guarracino (eds), 982–983. Milano: Bruno Mondadori.

Savoia, Leonardo M. 1995. Qualche riflessione sul linguaggio politico. In *Linguaggio e politica*, Carla Ciseri Montemagno (ed.), 19–26. Firenze: Le Monnier.

Savoia, Leonardo M. & Marotta, Giovanna. 1994. Vowel properties and nuclear constituents: Evidence from Italian dialects. *Probus* 6: 43–79.

Savoia, Leonardo M. 1994a. Parametrical variation in prosodic licensing: Evidence from Italo-Albanian dialects. *Rivista di Grammatica Generativa* 20: 105–141.

Savoia, Leonardo M. 1994b. Patologia del linguaggio. In *Enciclopedia Italiana – V Appendice*, 91–93, Roma: Istituto della Enciclopedia Italiana.

Savoia, Leonardo M. 1994c. Fonologia naturale. In *Enciclopedia Italiana – V Appendice*, 282–283. Roma: Istituto della Enciclopedia Italiana.

Savoia, Leonardo M. 1994d. Fonetica. In *Enciclopedia Italiana – V Appendice*, 281–282. Roma: Istituto della Enciclopedia Italiana.

Altimari, Franco & Savoia, Leonardo M. (eds). 1994. *I dialetti italo-albanesi*. Roma: Bulzoni.

Rizzi, Luigi & Savoia, Leonardo M. 1993. Conditions on /u/ propagation in southern Italian dialects: A locality parameter for phono-syntactic processes. In *Syntactic Theory and the Dialects of Italy*, Adriana Belletti (ed.), 252–318. Torino: Rosenberg & Sellier.

Savoia, Leonardo M. 1993. Condizioni fonologiche e morfosintattiche della coincidenza fra flessione media e attiva nei dialetti arbereshe di Barile e di Ginestra. *Quaderni del Dipartimento di Linguistica. Università degli Studi di Firenze* 4: 133–162.

Savoia, Leonardo M. 1992a. Educazione linguistica e diversità linguistica nella scuola italiana. Elementi per una riflessione. In *Scuola e società multiculturale. Elementi di analisi multidisciplinare*, G. Tassinari, G. Ceccatelli Gurrieri & M. Giusti (eds), 263–273, Firenze: La Nuova Italia.

Savoia, Leonardo M. 1992b. Acquisizione della fonologia in bambini italiani: aspetti teorici e questioni descrittive. In *Fonetica e fonologia nello sviluppo normale e patologico del linguaggio*, AA.VV. (eds), 65–82, Urbino: Università di Urbino.

Savoia, Leonardo M. 1992c. Coincidenza fra morfologia media e attiva nel dialetto di Barile. In *Dialetti italo-albanesi e letteratura*, Antonino Guzzetta (ed.), 61–70. Palermo: Istituto di Lingua e Letteratura Albanese.

Savoia, Leonardo M. & Marotta, Giovanna. 1991. Diffusione vocalica in un dialetto calabrese. Alcuni parametri fonologici. In *L'interfaccia fra fonologia e fonetica*, Elisabetta Magno Caldognetto & Paola Benincà (eds), 19–42, Padova: Unipress.

Savoia, Leonardo M. & Giannelli, Luciano 1991. Restrizioni sull'esito [*h*] da *t* in Fiorentino e nelle altre varietà toscane. *SILTA* 20: 3–57.

Savoia, Leonardo M. 1991. Alcuni elementi per una classificazione dei dialetti arbëreshë. In *La lingua albanese nell'Italia meridionale*. AA.VV. (eds), 13–52, Salerno: Università di Salerno.

Brandi, Luciana & Savoia, Leonardo M. 1990. Proprietà morfosintattiche e assegnazione del caso nel causativo arbëresh. *Rivista di Grammatica Generativa* 15: 29–121.

Savoia, Leonardo M. 1990a. Proprietà vocaliche e struttura metrico-sillabica in alcune varietà abruzzesi. In *Studi in memoria di Ernesto Giammarco*, AA.VV (eds), 331–380. Pisa: Giardini.

Savoia, Leonardo M. 1990b. Il caso della lingua italiana. In *IRRSAE Toscana. 39. Scuola Toscana (Bollettino dell'IRRSAE Toscana)* 1: 23–34,Firenze: IRRSAE Toscana.

Savoia, Leonardo M. 1989a. Processi morfologici, proprietà di caso e accordo nel causativo Arbëresh. *Zjarri* 33: 186–293.

Savoia, Leonardo M. 1989b. Alcune caratteristiche del causativo arbëresh. In *Le minoranze etniche e linguistiche*, AA.VV (eds), 363–420.

Savoia, Leonardo M. 1987a. Teoria generativa, modelli fonologici e dialettologia. La propagazione di /u/ in una varietà lucana. *Rivista Italiana di Dialettologia* 11: 185–263.

Savoia, Leonardo M. 1987b. Come gli adulti comunicano coi bambini. In *Gli italiani parlati. Sondaggi sopra la lingua di oggi*. AA.VV. (eds), 113–156. Firenze: Accademia della Crusca.

Savoia, Leonardo M. 1986a. La formazione di un modello linguistico 'neogrammaticale' nella linguistica italiana dell'ottocento. In *Un periodo di storia linguistica: I neogrammatici*, A. Quattordio Moreschini (ed.), 67–129. Pisa: Giardini.

Savoia, Leonardo M. 1986b. Sistemi flessionali e cambiamento morfologico: Affinità grammaticali e condizioni della variazione in alcuni dialetti arbëreshë. *Quaderni dell'istituto di linguistica dell'Università di Urbino* 4: 115–174.

Savoia, Leonardo M. 1985a. Per una educazione arbëresh. In *Le minoranze etniche e linguistiche*, AA.VV. (eds), 211–231, Palermo: Comune di Piana degli Albanesi – Facoltà di Lettere e Filosofia dell'Università di Palermo e Facoltà di Lettere e Filosofia dell'Università della Calabria.

Savoia, Leonardo M. 1985b. La morfologizzazione del vocalismo tematico nel perfetto in *o in alcune parlate arbëreshe. *Zjarri* XVII: 43–64.

Savoia, Leonardo M. 1984a. *Grammatica e pragmatica del linguaggio bambinesco (baby talk)*. Bologna: CLUEB.

Savoia, Leonardo M. 1984b. Appunti su alcuni aspetti del rapporto fra questione della lingua e pensiero linguistico. In *Lingua e dialetto*, Leonardo M. Savoia (ed.), 18–44, Pesaro: Comune di Pesaro.

Savoia, Leonardo M. 1984c. Rappresentazione delle differenze fra dialetti affini. *Quaderni dell'atlante lessicale toscano*, 2–3: 47–102.

Savoia, Leonardo M. 1983a. Ristrutturazione analogica: Sviluppi in alcune parlate albanesi. In *Scritti in onore di G. B. Pellegrini*, AA.VV. (eds), 1171–1187. Pisa: Pacini.

Savoia, Leonardo M. 1983b. Assetto della grammatica e realizzazione sociale nel cambiamento linguistico: Un esempio di variazione morfologica in una varietà arbëreshe. *Quaderni dell'istituto di linguistica dell'Università di Urbino* 1: 25–67.

Savoia, Leonardo M. 1981. Appunti per la storia della linguistica tra '700 e '800. In *Studi di linguistica italiana per Giovanni Nencioni*, AA.VV (eds), 351–420, Firenze: Il Pappagallo.

Savoia, Leonardo M. 1980a. Fonologia delle varietà apuane e garfagnine: Consonantismo. *Studi Urbinati di Storia, Filosofia e Letteratura. Supplemento Linguistico* 2: 233–293.

Savoia, Leonardo M. 1980b. La parlata albanese di S. Marzano di S. Giuseppe: Appunti fonologici e morfologici. *Zjarri* XII: 8–26.

Giannelli, Luciano & Savoia, Leonardo M. 1979–80. L'indebolimento consonantico in Toscana, II. *Rivista Italiana di Dialettologia* 3–4: 39–101.

Savoia, Leonardo M. 1979a. Aspetti dello sviluppo fonologico e morfofonologico del bambino: Studio di un caso. *Studi di Grammatica Italiana* VIII: 1–92.

Savoia, Leonardo M. 1979b. La trascrizione fonetica delle varietà lunigianesi: Appunti storici e metodologici. *Giornale Storico della Lunigiana e del Territorio Lucense* XXIX-XXX: 127–143.

Savoia, Leonardo M. 1979c. Review, A. Uguzzoni, *La fonologia*, Zanichelli, Bologna, 1978. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia*, III/IX: 2033–2038.

Giannelli, Luciano & Savoia, Leonardo M. 1978. L'indebolimento consonantico in Toscana, I. *Rivista Italiana di Dialettologia*, 2: 23–58.

Savoia, Leonardo M. 1978a. Lessico e fonologia del linguaggio bambinesco (baby talk) di area fiorentina. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia* VIII: 695–727.

Savoia, Leonardo M. 1978b. Review, Giuseppe Falcone *Profilo dei dialetti italiani, Calabria*, Pisa, Pacini 1976. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia* III/VIII: 1918–1921.

Savoia, Leonardo M. 1978c. Review, Patrizia Maffei Bellucci, *Profilo Dei Dialetti Italiani, Lunigiana*, Pisa, Pacini 1977. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia* III/VIII: 1914–1918.

Savoia, Leonardo M. 1975a. Prospettive teoriche e potenzialità di confronti nell'opera di Gustave Guillaume. In *Teoria e storia degli studi linguistici*, U. Vignuzzi (ed.), 361–389. Roma: Bulzoni,

Savoia, Leonardo M. 1975b. Condizioni fonetiche nel fiorentino comune e alcune proposte per una teoria fonetica concreta. *Studi di Grammatica Italiana* IV: 211–330.

Savoia, Leonardo M. 1975c. Review, A W. D'Addio Colosimo, *Lingua straniera e comunicazione. Problemi di glottodidattica*, Bologna Zanichelli, 1974. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia* III(V): 1754–1760.

Savoia, Leonardo M. 1975d. Review, I. Iordan & M. M. Manea, *Linguistica romanza*, Padova, Liviana, 1974. *Studi Mediolatini e Volgari* 23: 246–250.

Savoia, Leonardo M. 1974. Review, D. Parisi & F. Antinucci, *Elementi di Grammatica*, Torino, Boringhieri, 1973. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia*, III/LV: 1746–1754.

Micro- and macro-variation in syntax

Gender, number and inflectional class in some Northern Italian dialects

The plural inflection *-i* and the interpretation of N morphology

Benedetta Baldi and Ludovico Franco

University of Florence / New University of Lisbon

In many Northern Italian varieties the plural inflection *-i* has a restricted distribution, targeting feminine class nouns, determiners and adjectives. The *-i* morpheme also occurs with clitic pronouns, both subjects and objects. The crucial point we address is the relation between the plural specialized morpheme *-i* and its distribution among noun classes. We account for the distribution of the inflection *-i* on the basis of a theoretical treatment of the internal structure of the noun, along the lines of Manzini and Savoia (2017a, b).

Keywords: nominal class, gender, number, clitic, Northern Italian dialects

1. Introduction: The framework

In this paper we provide a morphosyntactic analysis of the plural morpheme *-i* in Northern Italian varieties, following the insight of Manzini and Savoia (2011a, 2017a, b). In recent work, Manzini and Savoia develop a morpheme-based approach to syntax, assuming that the same basic computational mechanisms underlie syntax and morphology.[1]

The morphemic analysis of Romance implies a first component which is a root, that, following Marantz (1997), we may conceive as category-less. Next to

---

1. Differently from other realizational approaches to morphology, in particular Distributed Morphology (DM) (Halle & Marantz 1993), in the model we pursue here the syntax projects from actual lexical items (morphemes), endowed with interpretable content (cf. also Kayne 2010). This excludes that a morphological component of grammar can rearrange and manipulate syntax prior to vocabulary insertion, as assumed in DM.

the root we find different kinds of morphemes, including a nominalizer element *n* (Marantz 1997), and derivational and inflectional ones. Inflectional morphemes generally follow derivational suffixes. A number of works (Kihm 2005; Ferrari Bridgers 2008; Picallo 2008), identify Marantz's (1997) *n* with the vocalic inflectional class/gender vowel immediately following the stem.[2]

The root is interpreted as a predicate (Higginbotham 1985) which has one open argument place (the R-role, Williams 1994) and is ultimately bound by a D/Q operator. Gender and number specifications apply to the argument, namely, they operate as predicates in turn, restricting the content of the argumental variable ultimately bound by a D(eterminer) or a Q(uantifier). Consider the representation in (1) for Italian *ragazz-o* 'boy' and *ragazz-a* 'girl'. The property 'ragazz-' is compatible with both a feminine and a masculine Class, depending on the sex denoted. We tentatively assign the inflectional vowel of Italian to an Infl Position, which embeds the root and the Class (Gender) nodes.

(1)

```
                        Infl
                       /    \
                 Class       Infl
                /     \      -o/-a
               √       Class_φ
             ragazz   [masc]/[fem]
```
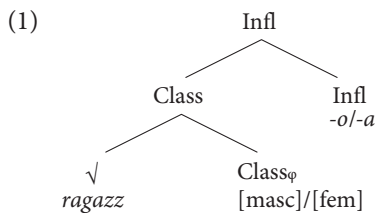
Italian.

In Manzini and Savoia's framework, the three inflectional class vowels in Italian, namely *-a, −o, −e* cover both masculine and feminine and partition the roots among them. Thus, for instance, √*madr* selects for *-e*, while √*mamm* selects for *-a*, in spite of sharing the same gender and the same meaning (mother).[3]

Consider now the plural. For the sake of clarity, let us take the structure of the plural nouns in Spanish, where a specialized *-s* inflection occurs. Given (1), if we identify the vocalic inflection of Spanish with the Infl position, the specialized *-s* segment for plurality in Spanish must occur on top of Infl itself, as schematized for

---

2. There is considerable consensus on the idea that the inner core of an N(P) is represented by a non-categorized root, immediately dominated by functional layers which embed the root into a nominal classification system (cf. Fábregas 2012; Kramer 2014). The properties that are directly relevant here are gender and countability (the latter often rendered as a Num category, cf. Ritter 1991; Picallo 2008; cf. also Borer 2005). The crucial split in the literature concerns whether the nominalization of the category-less root is prior to its classification (as seems to be assumed by Déchaine et al. 2014) or if it is the classification system that realizes *n* (Kramer 2014, 2015).

3. A different problem is the sexed interpretation of *ragazz-o/ragazz-a*. A possible answer is to assume that the sex interpretation of [fem] and [masc] is derived at the Conceptual-Intentional interface level, that is, it is a pragmatic effect due to knowledge of the world.

libr-o-s 'books' in (2). Following Manzini and Savoia (2011a, b) we formalize the content of the plural node as ⊆; this says that the denotatum of the predicate can be partitioned into subsets.

(2)
```
                        [⊆]
                       /    \
                    Infl     [⊆]
                   /    \       s
                Class    Infl
               /    \      o
              √     Class
            libr   [masc]
```

<div align="right">
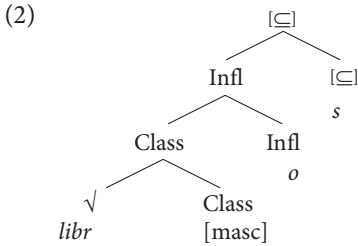
Spanish.

</div>

Differently from languages like Spanish, endowed with an independent lexicalization for the plural, namely *-s*, the plural in Italian is obtained by a change of the inflection. Thus, we may suppose that the plural of *ragazz-o/ragazz-a* in (1), namely *ragazz-i* 'boys', *ragazz-e* 'girls' has the structure in (3), where we assign the specification [⊆] to the morpheme *–i*, assuming that it is specialized for this interpretation.

(3)
```
                Infl
               /    \
            Class    Infl
           /    \    -e/ -i [⊆]
          √     Class
       ragazz  [fem]/[masc]
                  [⊆]
```

<div align="right">
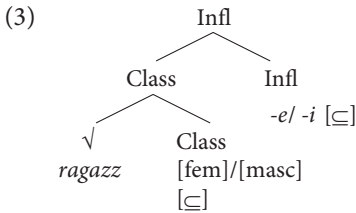
Italian.

</div>

Since we avoid inflectional class diacritics in such a model, it is important to note that plurals cannot be predicted from singulars in Italian. Thus, feminines inflected by *-a* can have an *-e* plural, as in (3), but also an *-i* plural (e.g. *al-i* 'wings') – vice versa *-i* plurals correspond to *-e, −a, −o* singulars. As to the nature of the exponent *-i*, Manzini and Savoia (2011a) connect its semantic value with its distribution in the Latin nominal paradigms. Their proposal is that (⊆) contributes plurality to the noun by isolating a subset of all individuals that are defined by the predicative content of the nominal root (cf. Chierchia 2010). For what concerns the *−a* inflection, we remind that it is able to introduce a special type of plurality in many Italian varieties, including the Standard one (Acquaviva 2008).[4] Manzini and Savoia (2017a, b) assume that it contributes to denoting a set whose members

---

4. Italian examples of lexical items taking both *–i* and *–a* plural morphemes include the pairs *oss-a/oss-i* 'bones', *muri/mur-a* 'walls', *dit-i/dit-a* 'fingers', where the *–a* inflection conveys a collective/aggregate interpretation.

are rather more like parts of whole than like individuated atoms, characterizing it as [aggregate]. The notion of an aggregate of parts is used by Chierchia (2010) to characterize mass denotation (cf. Franco et al. 2018). Assuming the existence in Romance (Indo-European) of an [aggr] class, it is tempting to differentiate the -*a* plural from the -*i* plural by associating with the former the properties [⊆,], [aggr].[5]

2.   The distribution of the plural -*i* in North Italian dialects

We begin by examining the Northern Italian dialects in which -*i* is an exponent for the feminine plural with nouns and adjectives (participles), and shows up with the determiners, in feminine and masculine contexts. In these systems, *(−)i* generally externalizes the plural also with subject and object clitics (see Section 3.1). All the data discussed here are taken from Manzini and Savoia (2005: § 4.6.2ff). This distribution of the -*i* morpheme specifically interests many varieties spoken in the Po Valley and shows a subtle micro-variation. In the South Lombard dialect of Revere, singular feminine nouns (and adjectives) present the inflection -*a* (4a) or lack any inflectional morpheme (4e). The plural is specified by the exponent -*i* only with the subclass presenting the inflection -*a* (4a′); the other feminines lacking inflectional morphemes, like *no:s* 'walnut/walnuts' are devoid of any plural exponent (4e′). With determiners, −*i* combines with the definiteness base *l*-, in (4a′, e′). The combination -*l-i* may be included also within the determiner *k-l-i* 'those.f' (4a′). Masculine nouns are generally devoid of inflectional material also in the plural forms and coincide with the root (4c). However, the plural determiner coincides with the morpheme *i*, but is found only with modifiers and a restricted subset of adjectives, like *bɛ-i* 'fine.M.PL'. In predicative contexts, the same distribution of the inflectional exponents is attested, as illustrated in (4b), (4b′) and (4d), (4d′).

---

5. Separating gender (and number) from inflectional class, has the consequence of recognizing the crucial role of Agree, since agreement between the noun, its modifiers, and it determiners is sensitive to gender and number (as well as case), but not to inflectional class. In the minimalist framework (Chomsky 2000, 2001), agreement processes are standardly associated with the rule of Agree – which however is conceived so as to account for one-to-one agreement in the sentential domain. In what follows, we will keep the assumption that Agree also applies within DPs. We will assume that what impels Agree to apply is the necessity of creating equivalence classes of phi-feature bundles denoting a single referent (the equivalent of uninterpretable feature deletion) (Manzini & Savoia 2005, 2007, 2011a, b). In other words, departing from current minimalist practice, we assume that there are only positive properties in language. Therefore, there are no uninterpretable properties; all lexical material is interpreted at the Conceptual-Intentional (C-I) interface. We can say that an argument agrees with the predicate, in the sense that the Identity Relation (Match) holds of them (cf. Franco et al. 2015).

(4) *Revere*

    a.   l-a   dɔn-*a*
         the-F.SG woman-F.SG

         nɔstr-*a*   fiøl-*a*
         our-F.SG daughter-F.SG

         kl-a   dɔn-*a*   la
         that-F.SG woman-F.SG   there

         st-a   dɔn-*a*   ki
         this-F.SG woman-F.SG here

         kl   altr-*a*   dɔn-*a*
         that   other-F.SG   woman-F.SG

         kl-a   bɛl-*a*   dɔn-*a*
         that-F.SG pretty-F.SG   woman-F.SG

   a′. l-i   dɔn-i
         the-PL   woman-F.PL

         l-i   nɔstr-i   fiøl-i
         the-PL   our-F.PL   daughter-F.PL

         kl-i   dɔn-i
         that-PL woman-F.PL

         st-i   dɔn-i
         this-PL woman-F.PL

         k-i   altr-i   dɔn-i
         that-PL other-F.PL   woman-F.PL

         k-i   altr-i   bɛl-i   dɔn-i
         that-PL other-f.p   pretty-F.PL   woman-F.PL

         pɔk-i / kwant-i   dɔn-i
         few-.PL/ many-F.PL   woman-F.PL

         kl-i   dɔn-i   vɛtʃ-i
         that-PL woman-F.PL old-F.PL

    b.   l   ɛ kwɛst-*a*   ki/ kwel-*a*   li
         she is this-F.SG   here/that-F.SG there

         l   ɛ bɛl-*a*/vɛtʃ-*a*
         she is pretty-F.SG/old-F.SG

   b′. i   ɛ   kwest-i ki/kwel-i   li
         they   are   this-PL here/that-PL there

         i   ɛ bɛl-i /   vɛtʃ-i
         they are pretty-F.PL/ old-F.PL

c.   l   ɔm
the   man
kɐl   kaŋ
that   dog
nɔstɐr fiˈøl
our   son
kl   ɔm
that   man
st   ɔm
this   man
kl   altr   ɔm
that   other   man
kɐl   braʊ/ bɛl ɔm
that good/nice man
kl   ɔm   vɛːtʃ
that   man   old

c′. i   ɔm
the-PL man
i   nɔstɐr   fiˈø-i
the-PL our   son-PL
k-i   ɔm
that-PL man
st-i   ɔm
this-PL man
k-i   altɐr   ɔm
that-PL other   man
k-i   braʊ/bɛ-i   ɔm
that-PL good/nice-PL man
k-i   ɔm   vɛːtʃ
that-PL man   old
pɔk-i / kwant-i   ɔm
few-PL/ many-PL   man

d. l ɛ kwestu ki / kwel   li
he is this.ONE here/that.ONE there
l ɛ ʋɛtʃ / bɛːl
he is old/nice

d′. i   ɛ kwest-i   ki/ kwel-i   li
they are this-PL   here/that-PL there
i   ɛ ʋɛːtʃ / bɛ-i
they are old/nice-M.PL

    e. l-a   no:s
        the-F.SG walnut
    e′. l-i   no:s
        the-PL walnut
        k-i   altr-i   no:s
        that-PL other-PL walnut
        k-i   altr-i   bɛl-i no:s
        that-PL other-PL nice-PL walnut
        kl-i   no:s   vɛtʃ-i
        that-PL walnut   old-PL

The data of Moncalvo (East Piedmont) in (5) show a similar pattern. Feminines distinguish between singular *-a*, in (5a), and plural *-i*, in (5a′); masculines lack any inflectional exponent, in (5c), (5c′), except for a subset of nouns/ adjectives that introduce *-i* in the plural, like *omn-i* 'men', *bɛ-i* 'fine.M.PL.', in (5c′). Masculine determiners and other modifiers include *(−)i* or realize an alternant with a palatalized final consonant. In the feminine plurals, two alternants of the article occur, *al* before words with an initial consonant and *i* before words with an initial vowel, as illustrated in (5a′) by the contrast between *al dɔni* 'the women' and *i atri dɔni* 'the other women'. The other modifiers present the inflection *-i*. It is of note that in some modifiers the masculine plural differentiates from the feminine one. More precisely, while the feminine inserts *-i*, like *pɔk-i* 'few.f.', *tant-i* 'many.f.', the masculine realizes *-i* as a palatalized outcome, like *tantʃ* 'any-m.PL', *pɔik* 'few-m.PL'. Again, there are also feminines (5e), (5e′) which lack any inflectional exponents and coincide with the root.

    (5) *Moncalvo*
        a. l-a   dɔn-a
            the-F.SG   woman-F.SG
            mɛ   surɛl-a
            my   sister.F.SG
            kust-a/ s-a/ kul-a   dɔn-a
             this.F.SG/his.F.SG/that.F.SG   woman-F.SG
            l   atra   dɔna
            the   other.F.SG woman.F.SG
            kula   bɛla   dɔna
            that.F.SG   pretty.F.SG woman.F.SG
            kula   dɔna   vɛdʒa
            that.F.SG   woman.F.SG old.F.SG

a′. al   dɔn-i
     the.PL   woman-F.PL
     al   mɛ   surɛl-i
     the.PL   my   sister-F.PL
     kust-i/   s-i /kul-i   dɔni
     this-PL/ their-PL/that-PL woman.F.PL
     i   atr-i   dɔn-i
     the.PL   other-PL   woman-F.PL
     kul-i   bɛl-i   dɔn-i
     that-PL   pretty-PL   woman-PL
     kul-i   dɔn-i   vɛdʒi
     that-PL   woman-F.PL   old.PL
     pɔk-i /   tant-i   dɔn-i
     few-PL/   many-PL   woman-F.PL

b.   al   e kul-a   li/kuçt-a
     she is that.ONE-F.SG there/this.ONE-F.SG
     al   ɛ vɛdʒ-a
     she is old-F.SG

b′. i   suŋ kul-i/ kuçt-i
     they are that-PL/ this-PL
     i   suŋ vɛdʒ-i
     they are old-PL
     i   suŋ dɔn-i / pɔk-i
     they are woman-F.PL/ few-PL

c.   l   ɔ:m
     the   man.SG
     al   ga:t
     the.M.SG   cat
     al   kurˈtɛ
     the.M.SG   knife
     kust/ is/ kul   ga:t
     this/his/that-M.SG cat
     kul   bɛl   ga:t
     that.M.SG   nice.M.SG cat
     kul   ga:t   vɛ:tʃ
     that-M.SG   cat   old

c′.   i   ɔmn-i
       the.PL   man-PL
       i   ga:t
       the.PL   cat
       i   kur'tɛ
       the.PL   knife
       kust-i/ si /ku-i   ga:t
       this/their/that-PL   cat
       ku-i   bɛi   ga:t
       that-M.PL   nice-M.PL cat
       ku-i   gat   vɛ:tʃ
       that-M.PL   cat   old
       pɔik/tantʃ   ga:t
       few/many   cat

d.   al ɛ kul   li /kust
       he is that.ONE there/this.ONE
       al ɛ vɛ:tʃ
       he is old

d′.   i   suŋ kui   li/kusti   kwi
       they are that.one.PL there/this.ONE.PL here
       i   suŋ vɛ:tʃ
       they are old
       i   suŋ pɔik/tantʃ
       they are few/many

e.   la   nu:s
       the.F.SG walnut

e′.   al   nu:s
       the.PL walnut

In some other Northern Lombard dialects no specialized plural inflection occurs with nouns. This is the case of the variety of Civate in (6). Again, feminine introduces –*a* in the singular (6a), (6b), but in the plural the bare root shows up (6a′), (6b′). Masculines are normally inflectionless, both in the singular (6c), (6d), and in the plural (6c′), (6d′). Determiners and the other modifiers include *(−)i*, like the article *i* 'the', the demonstrative *k-i* 'those', and the modifier *pok-i* 'few' (6a′), (6b′), (6c′), (6d′). The inflection -*i* appears with a subset of nouns, like *surɛ-i/ fradɛ-i* 'sisters/brothers', and in some adjectives like *brav-i/ bɛ-i* 'good.PL/fine.PL', exemplified in (5a′), (5b′), (5c′), (5d′).

(6) *Civate*

    a.   l-a   dɔn-a
        the-F.SG woman.-F.SG
        l-a   nɔst-a   surel-a
        the-F.SG our-F.SG   sister-F.SG
        kel-a   dɔn-a   la
        that-F.SG woman-F.SG there
        kel   oltr-a   dɔn-a
        that   other-F.SG   woman-F.SG
        kel-a   brav-a   dɔn-a
        that-F.SG good-F.SG   woman-F.SG

    a′. i   dɔn
        the-PL women
        i   nɔst   surɛ-i
        the.PL our   sisters-PL
        k-i   dɔn   la
        that-PL woman   there
        k-i   olter   dɔn
        that-PL other   woman
        k-i   brav-i   dɔn
        that-PL good-PL woman
        k-i   bɛ-i   dɔn   la
        that-PL pretty-PL woman   there
        pok-i   dɔn
        few-PL woman
        k-i   dɔn   vɛ:tʃ
        that-PL woman   old

    b.   l   ɛ kel-a   la / kest-a
        she is that-F.SG there/this-F.SG

    b′. i   ɛŋ k-i   la / kis   ke
        they are that-PL   there/ these here
        i   ɛŋ vɛ:tʃ
        they are old
        i   ɛŋ brav-i / bɛ-i
        they are good-PL/nice-PL
        i   ɛŋ pok-i
        they are few-PL
        i   ɛŋ dɔn
        they are woman

c.   l   ɔm
       the.SG   man
       ul   kaŋ
       the.SG   dog
       ul   nɔst   fradɛl
       the-M.SG   our   brother.M.SG
       kel   olter   omen   la
       that-M.SG other man   there

c′.   i   omen
       the.PL   men.PL
       i   nɔst   fradɛ-i
        the.M.PL   our   brother-M.PL
       k-i   olter   omen   la
       that-M.PL   other   men   there
       kw-i   bravi/ bɛi   omen
       that-M.PL   good.PL/nice.PL   men
       pok-i/ tant-i   omen
       few-M.PL/many-M.PL   men
       k-i   omen   vɛ:tʃ
       that-M.PL   men   old

d.   l   ɛ   m   brau   omen
       he is a   good   man
       l   ɛ   k-el   la / k-es   ke
       he is that-sg there/this-sg here

d′.   i   ɛn brav-i   omen
       they are good-PL   men
       i   ɛŋ k-i   la / k-is   ke
       they are that-PL   there/this-PL here
       i   ɛŋ vɛ:tʃ /nø:v
       they are old/new
       i   ɛŋ pok-i
       they are few-PL
       i   ɛŋ omen
       they are me

We sum up the data above in (7), illustrating the occurrence of *(−)i* in the varieties considered. The association of a given D is independent of the occurrence of *-i* in the plural of nouns; on the contrary the lack of the specialized plural inflection in the nouns may imply its realization in D, although this relation is not mandatory. The point is that in these dialects *-i* characterizes only (or preferably) feminine nouns, while masculines are generally devoid of any inflection.

(7) Presence of the *-i* plural inflection

	fem/ masc D	fem nouns	subset of masc nouns	clitics.m/F.PL
Revere	+	+	-[6]	+
Moncalvo	+	+	+	+
Civate	+	-[7]	+	+

We can wonder what is the relation between inflectional class, gender and plural. In fact, bearing the analysis proposed in Section 1 in mind, −*i* is associated with the specialized property [⊆], which specifies the normal plural, identifiable with the relation of inclusion (part-whole) (Chierchia 2010; Manzini & Savoia 2017a, b).[8]

3.   Plural and D field

Turning now to the link between D and the realization of the specialized plural morpheme, apparently comparable data come from other languages in which a subset of the inflectional properties are selected only on determiners or other modifiers of the noun. Costa and Figueiredo (2004) analyse some Brazilian Portuguese varieties in which the plural inflection only occurs on the determiners or prenominal adjectives, as in *O-s/est-es/algun-s/un-s livr-o muit-o bonit-o* 'The/these/some book very nice'.[9] In Brazilian Portuguese this restricted distribution of *-s* involves both feminine and masculine. A point made by Costa and Figueiredo is potentially relevant for us, namely the distinction between singleton and dissociated morphemes. In the approach of Embick and Noyer (2001), agreement and case morphemes are not syntactic projections and are not represented in syntax but they are added post-syntactically "during Morphology". Typically, dissociated morphemes convey an information "separated from the original locus of that information in the phrase marker" (Embick & Noyer 2001: 557). In this line, Costa

---

6. In Rovere we do not find a proper subset of plural masculine nouns in *-i*. It is found just with the item *fi'ø-i* (sons) in (4c′).

7. The only exception is the word *surɛ-i* (sisters). The plural of feminine is consistently found without any *-i* inflection.

8. More precisely, the morpheme *-i per se* is not specifically connected with gender, as shown by its distribution in feminine and masculine agreement contexts. A first conclusion concerning (7) is that the specialization of the inflection *-i* for feminine nouns in Revere and Moncalvo dialects, as illustrated in (4) and (5), can be explained as the result of a referential split, so that only a subset of nouns is defined in terms of the content specified by the inflectional morphemes.

9. This kind of plural encoding is quite widespread among natural languages, as documented in Dryer (2013: WALS, Chapter 33), who dubs this strategy as 'plural clitics/words'.

and Figueiredo (2004) treat the plural morphology in Brazilian Portuguese as a singleton, i.e. a specialized interpretable morpheme, which associates only with the "element anchoring the information concerning number", namely Determiners.[10] In European Portuguese, where the plural agreement occurs on all of the elements internal to DP, it is a dissociate morphemes "associate post-syntactically with all items able to bear plural mark […]; in Brazilian Portuguese it is a singleton, a morpheme that "anchors on the category relevant for the interpretational component" (pp. 24–26).

Actually, our data are less transparent, since *-i* can appear on a subset of nouns and adjectives, and only the data of Civate in (6) seem to be close to the ones of Brazilian Portuguese. Our data introduce further elements of complexity. In the case of Civate the occurrence of *(−)i* is the only specialized realization of plural. However, it includes many contexts, occurring either as inflection of the determiners and pre-nominal adjectives or as autonomous lexicalization of plural articles and subject/object clitics. Moreover, masculine nouns lack any other type of inflection; similarly, feminine plural nouns coincide with the root. As to the data in (4) and (5), the plural specialized morpheme *(−)i* combines with the nouns, even if only with the nouns with feminine gender belonging to the class with singular inflection *–a*. Besides, in Moncalvo dialect in (5) the plural of the article in the feminine contexts includes also the alternant *al*, i.e. a form devoid of *–i*.

As sketched in Section 1, we assume that morphemes convey an interpretable content. From this perspective, the distinction dissociated/singleton could be rephrased as the split between agreement properties with general occurrence on the one side and those associated only with one category on the other side, i.e. with specialized occurrence. In the varieties in (4)–(6) the morpheme *(−)i* could be understood as a specialized exponent, seeing that it realizes the plural in determiners/ modifiers, in subsets of nouns (in some dialects) and is also found with subject or object clitic, as we will show in the next section. Once more, in these varieties, as outlined above, the plural is not systematically externalized in nouns/adjectives; it remains a specialized property of a subset of lexical elements.

3.1   The inflection of clitics

Before going any further in the discussion, let us consider the occurrence of *(−)i* in the clitic system of these dialects, that are endowed with both subject and object clitics. (8) illustrates the subject clitics paradigm. Also in the case of subject clitics

---

10. Costa and Figueiredo (2004) reach this conclusion by differentiating the Spec-head configuration, that triggers the Subject-I agreement, from the D-N relation, where the plural singleton occurs.

paradigms, there is strong variation (Manzini & Savoia 2005). In (8) we present subject clitics occurring with lexical verbs and we can note that: i) in the singular, 3rd person clitics have two different specialized forms, *al* for the masculine and *l-a* for the feminine; ii) in the plural (3rd person), we find only the form *i*, or in Revere dialect *i* for the masculine and *l-i* for the feminine (8a). For what concerns subject clitics occurring with auxiliaries (8′): i) in the 3rd person singular we find the simple definiteness base *l*; ii) *-i* is found with 3rd person plural or, in Moncalvo, the clitic *i* occur also with the other plural persons. Namely, *−i* keeps being the unique plural form lexicalizing properties, in principle connected to subsets of referents.

(8)   *Revere*

a.   a   dɔrmi
      I   sleep
      at   dɔrmi
      you   sleep
      al/la   dɔram
      he/she   sleep.3sɢ
      a   durmema
      we   sleep.1ᴘʟ
      a   dur'mi
      you.ᴘʟ   sleep
      i/li   dɔram
      they.m/f sleep.3ᴘʟ

a′.   u   dur'mi
      have.1sɢ slept
      t   ɛ   dur'mi
      you   have.2sɢ slept
      l   a   dur'mi
      he/she   has   slept
      e:m   dur'mi
      have.1ᴘʟ slept
      i   dur'mi
      have.2ᴘʟ slept
      i   aŋ   dur'mi
      they   have-3ᴘʟ slept

b. *Civate*
dɔrmi
sleep.1sg
te   dɔrmɛt
you   sleep.2sg
al/ la   dɔrma
he/she sleep.3sg
dɔrmɛm
sleep.1pl
dɔrmɛf
sleep.2pl
i   dɔrmɛn
they   sleep.3pl

b′. o   dur'mi
have.1sg slept
t   e   dur'mi
you   have.2sg slept
l   a   dur'mi
he/she has   slept
em   dur'mi
have.1pl slept
i   dur'mi
have.2pl slept
i   aŋ   dur'mi
they   have.3pl slept

c. *Moncalvo*
drɔ:m
sleep.1sg
ad   drɔmi
you   sleep.2sg
al/la   drɔ:m
he/she sleep.3sg
durmuma
sleep.1pl
a   drɔmi
you.pl sleep.2pl
i   drɔmu
they   sleep.3pl

c′. i   ɔ   dru'mi
     I   have.1sɢ slept
     t   ai   dru'mi
     you   have.2sɢ slept
     l   a   dru'mi
     he/she has   slept
     i   uma   dru'mi
     we   have.1ᴘʟ slept
     i   ei   dru'mi
     you   have.2ᴘʟ slept
     i   aŋ   dru'mi
     they   have.3ᴘʟ slept

In (9)–(11), we illustrate object clitics and participial agreement. In (9a)-(9a′) the data concerning the 3rd person forms is given. In the singular normally two different clitics for the masculine and feminine are attested. In the plural, two types of externalization are found: the plural is differentiated between feminine and masculine, as *i-a* 'them.m' and *l-i* 'them.f' in Revere in (9a), otherwise a common form occurs, like *j-i* (Moncalvo), (10a), or *i-a* (Civate), (11a). As investigated in Manzini and Savoia (2005: § 4.6.2ff.), in these dialects a constraint on the inflectional properties of the 3rd person object clitics generally emerges in contexts in which they combine with 3rd person subject clitics, as in (9a′)-(11a′). So, in (9a′), the 3rd person object clitic is *l-a* for both gender classes when preceded by the 3rd person masculine subject clitic. In the case of Civate, in (10a′), in the contexts of 3rd person, the subject clitic is not lexicalized and the object clitic presents the only form *l-a* for masculine and feminine singular and *i-a* for the plural. There are dialects that do not present this constraint as the one of Moncalvo, where two distinct clitics occur also in these contexts (11a). Finally, the participles in (9b)-(11b) show the same inflection seen within DPs. Crucially, it is the inflection of the participles that registers the gender class properties of the clitics *l,* as the *–i* inflected past participle agreeing with 3rd person feminine clitic in Revere in (9b). In the Piedmontese dialect of Moncalvo the clitics attach to the uninflected form of the participle, doubling the proclitic form, as exemplified in (11b).

(9) a. at   al   / l-a / i-a / l-i   tʃami   *Revere.*
         you him/ her/ them / them.ꜰ.ᴘʟ call.2sɢ
     a′. ly: al   l-a   / i-a tʃama /
         he ᴄʟs.ᴍ him/ her / them he.call.3sɢ
         le: la   l / l-a   / i-a tʃama
         she ᴄʟs.ꜰ him/ her / them call.3sɢ

 b. a l em   tʃamad-*a* /
 CLS her have.1PL called-F.SG
 a i em   tʃamad-i /
 CLS them have.1PL called-F.PL
 a l em   tʃa'ma /
 CLS him have.1PL called.M.SG
 a i em   tʃa'ma
 ClS them have.1PL called-F.SG

(10) a. al / l-a / i-a tʃamɛf   *Civate.*
 him/ her / them call.2PL
 a′. ly/ le l-a   tʃama/
 he/ she him/ her call.3SG
 l-a / i-a tʃamɛn
 him/ her / them call.3PL
 b. l o   tʃamad-a /
 her have.1SG called-F.SG
 i o   tʃama:/
 them have.1SG called
 l o   tʃama:
 him have.1SG called

(11) a. kil a l-u / l-a / j-i tʃama   *Moncalvo.*
 he CLS him/ her / them call-3SG
 b. (al) l   aŋ   tʃa'ma-l-u/l-a /
 CLS him/her have.3PL called-him/her
 (a) i   aŋ   tʃa'ma-ji
 CLS them   have.3PL called-them

In the singular, the feminine clitic has the form *l-a*, as shown in (9)–(11); on the contrary, in many Northern dialects the masculine singular corresponds to the *-l* morpheme only, possibly preceded by a vocalic element *e-* or *a-*, as in Revere and Civate. Interestingly, the feminine plural object clitic of Revere has a form richer than the masculine plural, including the definiteness base *l-*. The dialect of Moncalvo registers the singular masculine by means the morpheme *-u*.

4.   Some proposals for the analysis

The data so far examined evidence some main patterns:

1. *(−)i* plural morpheme occurs on determiners and modifiers of the noun
2. *(−)i* plural morpheme is admitted with feminine nouns (mainly the one which present an *(−)a*)
3. *(−)i* plural morpheme is lexicalized with both subject and object 3rd person plural clitics
4. Masculine nouns are, generally, devoid of any inflection

All in all, a generalization can be attained: whereby if one lexical form is richer, this is the feminine. (12) formalizes a sort of scale of referentiality, in terms of the preservation of a more specialized morphology with nominal items. In other words, the more an inflection has denotational strength, the more systematically it is retained.

(12)   *feminine singular > feminine plural > masculine*

We know indeed that the inflectional elements on a par with the other morphemes in the word contribute to introducing properties that contribute to fixing the reference of the noun. So, −*a* is normally retained in Northern Italian dialect to express the feminine singular, while the *i*- exponent for plural is less systematically retained, and, finally, masculine tends to exclude any inflectional exponent.

We can connect (12) with the split between feminine and masculine already suggested in the discussion in Section 2, whereby the specialization of the inflection -*i* for feminine plural nouns in Revere and Moncalvo dialects in (4) and (5) can be understood as the result of a referential split.[11]

If we are on the right track, one can expect that the interpretive salience of the feminine has an effect on its inflectional properties. In the systems we analyse, the specialized morpheme -*i* externalizes the plural interpretation in the subset of the feminine nouns. These, in Northern Italian varieties like in the majority of Romance varieties, preserve also the singular inflection -*a*. As we saw in Section 1, we have assumed that the Infl element -*i* is associated with interpretive content, namely [⊆]. As to the inflection -*a*, according to Manzini and Savoia (2017, a, b),

----

11. In the DM literature concerning Romance languages, a traditional view considers the -*o* inflection as a default devoid of gender specification (Harris 1991; Halle & Marantz 1994; Ferrari-Bridgers 2008). On the contrary, feminine is dealt with as a true gender element, crucially connected with the sexed interpretation. We find the default analysis untenable both on principled bases and on the basis of its regular connection with interpretive content. Yet, we derive from these proposals the insight that [fem] is a lexical specification endowed with an especially recognizable and salient interpretive content.

it is associated with the content [aggr]. If there is a single Infl item *-a* occurring in the (feminine) singular as well, [aggr] must be associated with it, although the [aggr] content manifests itself optionally, specifically in correspondence of some sort of plurality. In the absence of other restrictions, we predict nevertheless that the property [aggr] may be present on *-a* in the singular as well (cf. Nussbaum 2014 for historical evidence).

This conclusion is confirmed also by the forms *i-a* for the plural of the object clitics in (9a) and (10a). As one can see, this type of plural combines the morpheme *i-* specialized for [⊆] with the morpheme *-a*. We can conclude that it is the specification [aggr] that allows *–a* to contribute lexicalizing the plural reading. Indeed, given the discussion that precedes, we can point to a positively specified property of *-a* that bridges between singular and plural, namely [aggr]. In other words, it is by virtue of the property [aggr] that *-a* turns up both as a plural, and a singular inflectional class marker. The fact that *-a* happens to be something like 'feminine' is a matter to be learned by the child.[12]

Masculines are treated differently. Usually in Romance Italian varieties the masculine is associated to an individual interpretation. In these dialects masculines are devoid of the inflection and acquire a singular or a plural reading by means of binding by D.

Finally, the special semantic status of the inflection *–a* is highlighted by the insertion of the object clitic *l-a* instead of the masculine *al* in the contexts with a 3rd person subject clitic, as exemplified in (9a′) and (10a′). The insertion of the form *l-a* can go together with the lack of the subject clitic, as in (10a′) for Civate. Manzini and Savoia (2005) conclude that in many varieties the insertion of a 3rd person object clitic is able to interpret the whole of the denotational properties relevant both for the subject and the object. *l-* lexicalizes definiteness properties (Manzini & Savoia 2005). As to the inflection *-a*, it has a stronger denotational power, which is necessary in the contexts of reciprocal exclusion between subject and object clitics. We can make the denotational power of *-a* more precise, at this point. In other words, the property that makes *-a* suitable for externalizing specialized interpretive properties, is its connection with the denotational import that we have associated with the feminine Gender in terms of the quantificational specification [aggregate].

---

12. It is worth noting that the classical historical account of Indo-European feminine singular/ neuter plural *-a* (Clackson 2007: 107) is that a neuter/collective plural *-a* was extended to a new inflectional class for collective/abstract singulars – which only secondarily came to coincide with the default class for feminine animates. Viewed as a projection on the historical, external axis of an analysis motivated on internal grounds, this reconstruction appears to be quite compatible with the present discussion.

4.1   The nature of -i

The verbal paradigms in (8) confirm the idea that we cannot identify -*i* with a sort of an independent morphological element endowed with interpretable content, as maintained for the -*s* plural in Brazilian Portuguese by Costa and Figueiredo (2004). In fact, in the paradigms in (8) the 3rd person plural of the subject clitic externalize -*i* as a unique plural inflection in Moncalvo and Civate, and, at least for the auxiliary, in Revere as well. Following the generalization of Vikner (1997), the presence of a unique specialized inflection at least in a sub-set of forms allows us to assume that the usual verbal agreement mechanism applies, in terms of V-to-I and features checking traditional scheme. So, we conclude that subject clitics do the same work as determiners in contributing to lexicalizing the plural interpretation [⊆], as in (13). In other words, the subject clitics and articles in (13) lexicalize the plural content, independently of whether the verbal or nominal head have a specialized morphology for it.

(13) a. lura:tʃ/luratr-i i   drɔm-u   *Moncalvo.*
          they.M/F   CLS sleep.3PL
          'they sleep'.

      b. la   kl-i dɔn-i l-i   m tʃam-*a*   *Revere.*
          there those women CLS.PL me call.3PL
          'Those women call me'.

      c. i   ryɛn   i baga-i   *Civate.*
          CLS arrive.3PL the boys

'The boys arrive'

      The fact that the plural *(–)i* systematically lexicalizes the plural independently of gender distinctions means that its only content is the pure property of plural, i.e. the part-whole/sub-set relation [⊆]. This reading includes both the inflectional occurrence of -*i* and its occurrence as an object or subject clitic in (13), as in the representation in (14) for subject clitics.

(14)   ClP
         ╱   ╲
       Cl   IP
     i [⊆]   ╱   ╲
             I
           ╱   ╲
          √   I
       drɔm -   u[⊆]
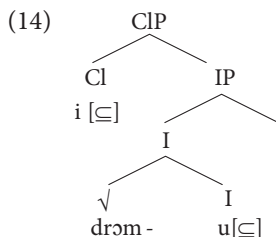
                                                                           Moncalvo.

Let's address now the question concerning why determiners take the plural *-i* inflection in all contexts. Costa and Figueiredo (2004) explain the association of the plural *-s* with determiners by assuming that "it must be attached to the head responsible for establishing the link with semantic interpretation", that is the determiner, introducing the definite reading. Naturally, we agree with the idea that determiners have a crucial interpretive role, in that they provide definiteness and other deictic informations to the C-I system. Manzini and Savoia (2005: § 8.2.3, 8.2.5), in discussing the distribution of the plural inflection *(−)i* in DP of Northern Italian varieties, conclude that the plural inflection "appears on the functional elements of the phrase in that they lexicalize the definiteness properties associated to the D position of the DP". In other words, we may expect some type of morpho-syntactic split, whereby definiteness and deictic elements are endowed with specialized morphology given the role that they play in the identification of the participants in (arguments of) the event. In the case at hand this is true for the *(−)i* morphology introducing the plural reading [⊆]. We assume that the relevant property is simply intrinsic to the lexical entries involved, namely determiners and clitics, and interpreted when merged in a sentential context.

4.2   The nominal inflection

Let's return now to the paradigms in (4)–(6). On the basis of the preceding discussion, we assign the plural *dɔn-i* (Revere) the structure in (15). More precisely, differently from the analysis provided for *ragazz-i/−e* in (4), the quantificational property [⊆] is associated only to the inflectional morpheme that selects the combination *root + [$_{class}$ fem].*[13]

(15)

Infl
├── Class
│   ├── √ (*dɔn*)
│   └── Class [fem]
└── Infl
    -i
    [⊆]
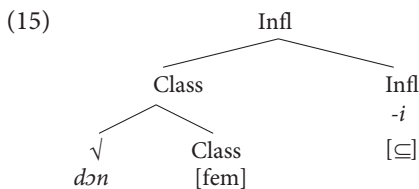
Revere.

In the case of masculines, generally no inflectional morpheme combines with the root. So, we assign the noun a structure lacking the inflectional node and coinciding with the bare Gender class, as in (16). What (16) highlights is the simplex nature of these lexical entries, as devoid of the number content. In other words,

---

13. The singular, in turn, is selected by the *-a* inflection, where – as outlined above – it is possible to assume that *-a* preserves its whole content, including the specification [aggr] (see Franco et al. 2018).

no number inflection can combine with and select these roots, thus attributing a mere predicative nature to them. Only D elements introduce the quantificational properties necessary for identifying the argument of the noun.

(16)

```
              D
           ／    ＼
         D         Class
         i_y      ／    ＼
        [⊆]     √        Class
               ɔm_y     [masc]
```
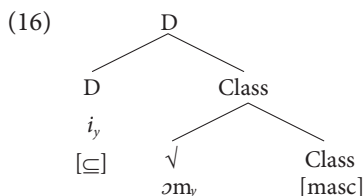
Revere.

In the dialect of Civate in (6) the plural feminine lacks in turn any inflection, coinciding with the bare root. Again we must conclude that it is the D operator that introduces the quantificational properties contributing to identifying the argument of the noun, as in (17).

(17)

```
              D
           ／    ＼
         D         Class
         i_y      ／    ＼
        [⊆]     √        Class
               dɔn_y    [fem]
```
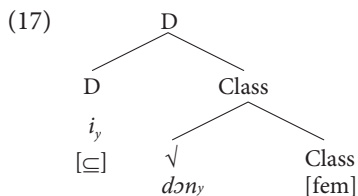
Civate.

In short, following Manzini and Savoia (2017a, b), we have proposed that inflections select the roots allowing for their Gender class and, possibly, other classificatory properties. In the case of nouns devoid of inflectional elements, we can conclude that the Gender is anyway associated to the root, as manifested by the agreement internal to DP and with the verb. Yet, it will be the determiners and the modifiers and in general the agreement contexts that provide for the interpretation with referential import. In the case of the plural the bare roots are lacking in number properties and wait for the determiners and modifiers for fixing the reference.

5.   Conclusions

In this work, we have compared some Northern Italian Romance varieties where the plural inflection *-i* occurs only in feminine nouns, in adjectives and in determiners. In these varieties only the inflection *-a* of the first class nouns (feminines) is preserved, while the other Gender classes lack any vocalic inflection and coincide with bare roots. The specialized plural inflection *(−)i* appears also in clitic pronouns, both subjects and objects, triggering feminine and masculine agreement.

We have addressed the relation between the plural specialized morpheme *(−)* *i* and its distribution in the noun classes and in the subject and object clitic system. Our proposal, grounded on the ability of the inflectional elements to convey specialized semantic content accounts for the morphological variation within a restrictive (minimalist) morpho-syntactic model. In the terms of this model of inflection, we have proposed that in the relevant Romance varieties, the semantic content of *-i* and *-a* can contribute to explaining their behaviour in nominal and sentential contexts.

Acknowledgements

References

Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: OUP.

Borer, Hagit. 2005. *Structuring Sense, Vol. 1: In name only*. Oxford: OUP.

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174: 99–149. https://doi.org/10.1007/s11229-009-9686-6

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step, Essays on Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels & Juan Uriagereka (eds), 89–155. Cambridge MA: The MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, Michael J. Kenstowicz (ed.), 1–54. Cambridge MA: The MIT Press.

Clackson, James. 2007. *Indo-European Linguistics: An Introduction*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511808616

Costa, João & Figueiredo Silva, Maria Cristina. 2004. Notes on nominal and verbal agreement in Portuguese, *Rivista di Grammatica Generativa* 27: 17–29.

Déchaine, Rose-Marie, Girard, Raphaël, Mudzingwa, Calisto & Wiltschko, Martina. 2014. The internal syntax of Shona class prefixes. *Language Sciences* 43: 18–46. https://doi.org/10.1016/j.langsci.2013.10.008

Dryer, Matthew S. 2013. Coding of nominal plurality. In *The World Atlas of Language Structures Online*. Matthew S. Dryer & Martin Haspelmath (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/33> (7 August 2017).

Embick, David & Rolf, Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4): 555–595. https://doi.org/10.1162/002438901753373005

Fábregas, Antonio. 2012. Structural and lexical aspects of the grammar of desinences. *Poznań Studies in Contemporary Linguistics* 48: 239–284. https://doi.org/10.1515/psicl-2012-0012

Ferrari Bridgers, Franca. 2008. A unified syntactic analysis of Italian and Luganda nouns. In *The Bantu-Romance Connection* [Linguistik Aktuell/Linguistics Today 131] Cécile De Cat &

Katherine Demuth (eds), 239–260. Amsterdam: John Benjamins.
https://doi.org/10.1075/la.131.13fer

Franco, Ludovico, Manzini, M. Rita & Savoia, Leonardo M. 2015. N class and its interpretation: The neuter in Central Italian varieties and its implications. Special issue on *Italo-Romance Morphosyntax*, Diego Pescarini & Silvia Rossi (eds). *Isogloss*, 41–68.

Franco, Ludovico, Baldi, Benedetta & Savoia, Leonardo M. 2018. Collectivizers in Italian (and beyond). *The interplay between collectivizing and evaluating morphology (and the Div paradox)*. Ms, Lisboa & Firenze.

Halle, Morris & Alec, Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20. Essays in Linguistics in honor of Sylvain Bromberger*, Kenneth Hale & Samual J. Keyser (eds), 111–176. Cambridge MA: The MIT Press.

Harris, James. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27–62.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16: 547–621.

Kayne, Richard. 2010. *Comparisons and Contrasts*. Oxford: OUP.

Kihm, Alain. 2005. Noun class, gender, and the lexicon/syntax/morphology interfaces: A comparative study of Niger-Congo and Romance languages. In *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque & Richard Kayne (eds), 459–512. Oxford: OUP.

Kramer, Ruth. 2014. Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Language Sciences* 43: 102–115.
https://doi.org/10.1016/j.langsci.2013.10.004

Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199679935.001.0001

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo M. 2007. *A Unification of Morphology and Syntax. Investigations into Romance and Albanian dialects*. London: Routledge.

Manzini, M. Rita & Savoia, Leonardo M. 2011a. *Grammatical Categories*. Cambridge: CUP.
https://doi.org/10.1017/CBO9780511974489

Manzini, M. Rita & Savoia, Leonardo M. 2011b. Reducing 'case' to denotational primitives: Nominal inflections in Albanian. *Linguistic Variation* 11: 76–120.
https://doi.org/10.1075/lv.11.1.03man

Manzini, M. Rita & Savoia, Leonardo M. 2012. Case categories in the Geg Albanian variety of Shkoder. *Res Albanicae* 1: 23–42.

Manzini, M. Rita & Savoia, Leonardo M. 2017a. N morphology and its interpretation: The neuter in Italian and Albanian varieties. In *Constraints on Structure and Derivation in Syntax, Phonology and Morphology*, Anna Bloch-Rozmej & Anna Bondaruk (eds), 213–236. Frankfurt: Peter Lang.

Manzini, M. Rita & Savoia, Leonardo M. 2017b. Gender, number and inflectional class in Romance. In *Proceedings of the Olomouc Linguistics Colloquium, 'Categorial features of Nouns and their projections'*.

Marantz, Alec 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4: 201–225.

Nussbaum, Alan J. 2014. Feminine, abstract, collective, neuter plural: Some remarks on each (expanded handout). In *Studies on the Collective and Feminine in Indo-European from a Diachronic and Typological Perspective*, Sergio Neri & Roland Schuhmann (eds), 273–306. Leiden: Brill. https://doi.org/10.1163/9789004264953_012

Picallo, Carme. 2008. Gender and number in Romance. *Lingue e Linguaggio* VII: 47–66.

Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In *Syntax and Semantics 25: Perspectives on Phrase Structure*, Susan Rothstein (ed.), 37–62. New York NY: Academic Press.

Vikner, Sten. 1997. V-to-I movement and inflection for person in all tenses. *Working papers in Scandinavian Syntax* 55: 1–27.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge MA: The MIT Press.

Objects and subjects in the left periphery

The case of a-Topics

Adriana Belletti

University of Siena & University of Geneva

A parallel is highlighted between the linguistic behavior of young Italian-speaking children as emerged in recent experimental work (Belletti and Manetti 2017) with comparative data from other Romance languages such as Balearic Catalan and (previous stages of) Spanish in the domain of a-Topics. A possible path from a-Topics to DOM is drawn, building on previous proposals in Belletti (2017a). a-Marking of topics concerns direct object DPs. A proposal is sketched out as to why subject DPs could not be a-marked when they are also topics in terms of a possibly conflicting requirement arising from the simultaneous satisfaction of both the Topic criterion with a-marking and the Subject criterion (in the sense of Rizzi 2006).

Keywords: a-Topics, left periphery, vP-periphery, development

1. Introduction

In (non southern, central variety of) Italian, which is not a Differential Object Marking/DOM language, full lexical direct object DPs may sometimes be introduced by preposition 'a' when they are left dislocated topics. This is, however, a severely constrained and limited possibility, only very marginally allowed by many speakers. The most felicitous cases involve left dislocated object experiencers of psych verbs (of the *worry* class, Benincà 1986; Belletti & Rizzi 1988; Berretta 1989; Belletti 2017a for recent discussion). a-Marked left dislocated topics are instead felicitously allowed by (non southern, central variety) speakers of Italian when the left dislocated object is a personal (strong) pronoun. Experimental results have brought to light the fact that developing children speaking the same variety of Italian appear to make a larger use of a-marking of left dislocated direct object topics by overwhelmingly a-marking left dislocated direct objects which are full lexical noun phrases and which are not object experiencers. Hence children

have overextended a limited possibility of adult Italian. After reviewing the main properties of children's a-Topics, crosslinguistic and diachronic parallels with the Italian speaking children's somewhat creative linguistic behavior are illustrated in Section 2.[1] One main distributional property that generally characterizes a-Topics is further discussed: they are objects they are not subjects. To the extent that this is a correct descriptive generalization, the question that arises is why it should be so. A possible proposal is sketched out in Section 3.

2.   a-Topics

2.1   a-Topics in Italian during development

Recent experimental findings on the acquisition of Italian presented and discussed in detail in Belletti & Manetti (2018) have revealed that when the production of a ClLD is elicited in which the direct object of a transitive verb occupies a left dislocated peripheral topic position and an accusative resumptive clitic is present in the sentence following it, young (4 to 6) Italian speaking children tend to realize the lexical direct object in the form of an a-Topic as illustrated in (1) :

(1)   Il   coniglio a i'   pinguino lo   tocca
      the rabbit   to the penguin him.CL touches
      'The rabbit is touching the penguin'
                              (Adele 4;9 – *Picture described: Rabbit touching penguin*)

Note that in (1) not only the direct object has been left dislocated, the subject as well is in a left peripheral topic position as it precedes the left dislocated direct object yielding the order SOClV. Italian is a language that allows for multiple topics (contrary to e.g. English, Rizzi 2013) and children appear to master this option quite well at the age investigated.

     In the references quoted (footnote 1) it is argued in some detail that the left peripheral direct object a-Topic of productions like (1) is not the manifestation of the Differential Object Marking/DOM phenomenology as direct objects in clause internal object position are not a-marked in the Tuscan variety of Italian that the tested children spoke. And indeed this is confirmed by the fact that the same children who produced a substantial amount of ClLDs of the type in (1) never a-marked the direct object when it had not been left dislocated yielding the production of a simple SVO declarative. The production in (2) , reproduced from the quoted references, illustrates this robust fact:

---

1. Belletti & Manetti (2018), (Belletti 2017a, b) for detailed description and discussion.

(2) La giraffa sta leccando <u>la mucca</u>, e   il   coniglio <u>al pinguino</u>   lo
     the giraffe is licking   <u>the cow</u> and the rabbit   <u>to the penguin</u> him-CL
     sta grattando.
     is scratching
     'The giraffe is licking the cow and the penguin is scratching the rabbit.'

(Omar, 5)

The fact that a-marking of direct objects is a left peripheral phenomenon affecting left dislocated direct object topics is not a weird and isolated fact concerning Italian speaking children during their development. Rather, children's behavior is robustly attested both crosslinguistically and diachronically. As it is often the case in acquisition, during their development children typically adopt grammatical options that are manifested in the grammar of some existing language.[2]

2.2   a-Topics in a comparative perspective

For instance, according to the insightful discussion in Leonetti (2004 and related work), modern Spanish is a DOM language in which direct objects are generally a-marked when they fill the clause internal object position (and are definite, specific etc.).[3] In some cases, however, depending on both the nature of the object (typically an indefinite) and the type of verb, a-marking may be optional when the object is in clause internal position. It becomes obligatory though if the same type of direct object is left dislocated in a ClLD structure. Hence, clitic left dislocated

---

2. This 'grammatically creative' behavior is typically due to some pressure conditioning children's immature grammatical system. The point is addressed in the references quoted where the pressure is identified in this case with the difficulty in handling an intervention computation in which the lexical direct object is moved across an intervening lexical subject along the lines of the system developed in Friedmann et al. (2009) in terms of featural Relativized Minimality (Rizzi 1990, 2004). See Belletti (2017)b for detailed discussion of this point and Belletti & Manetti (2018) for a general discussion along the same lines. This aspect will not be further addressed here.

It is also worth pointing out that children also used a-marked topics in non-intervention configurations and often with a generic plural null subject, as discussed in Belletti & Manetti (2018). This indicates that a-marking is not simply a strategy to disambiguate between the subject or the object interpretation of the two DPs.

3. The interesting literature on DOM that tries to identify the crucial features of the phenomenon is quite extensive, especially in the typological tradition (Bossong (1991 and much subsequent work); Iemmolo (2010) for discussion and an overview, and Manzini & Franco (2016) for a novel assessment. No attempt will be made here to provide an exhaustive characterization of the articulated phenomenology for both space reason and in the interest of a focused presentation. Only those crucial and robust patterns that are found both cross-linguistically and diachronically will be considered as the leading descriptive factors in the discussion in the text.

lexical direct object topics are always and obligatorily a-Topics in Spanish (Belletti 2017a). The same strong requirement on a-marking of left dislocated direct object topics is found crosslinguistically in various other (Romance) languages manifesting DOM to some extent, as is the case for example of the Sicilian variety described in Iemmolo (2007, 2010).

   The case that most closely resembles the described behavior displayed by Italian speaking developing children is provided by the Balearic variety of Catalan. According to Escandell-Vidal (2009) Balearic Catalan is not a productive DOM language. In particular lexical direct objects, differently from pronominal ones, are not a-marked when they fill the clause internal direct object position. However, they are typically, though not strictly obligatorily, a-marked when they are left dislocated.[4] The following examples from Escandell-Vidal (2009) illustrate the difference between SVO declaratives, with no a-marked object (3) , and the a-marked object in ClLD (4) :

(3) Balearic   (Arxiu, Maó. Minorca)
   a. He   anát a agafar es   qui ha   vist com
       have.PRES.1SG gone to catch the (one) who have.PRES.3SG seen how
       matava   en Fulano o en Sutano
       kill.PST.3SG the Fulano or the Sutano
       'I went to catch the one who has seen how he killed the Fulano or the
       Sutano'
   b. I   va   anar ja   a amenaçar es general
       and have.PST.3.SG go   already to menace   the general
       'So he went to threaten the general'   (Escandell-Vidal, 2009: 24a, b)

(4) Balearic   (Arxiu, Sant Josep, Ibiza)
   a. An aquella al.LOTA no la deixaven parlar amb so que ella
       to that   girl   not her let.PST.3PL talk   with the who she
       volia
       want.PST.3SG
       'That girl was not allowed to talk to the one she wanted'
                                                     (Escandell-Vidal, 2009: 27)

---

4. And also when they are right dislocated. For space reasons I will concentrate my discussion on left dislocated direct objects, which are also those robustly present in the children's productions of the Belletti & Manetti 's (2018) experiment.

b.   Balearic   (Majorcan)
    An aquesta darrera [frase]   noltros la diríem   així
    to this   last   [sentence] we   it sayCOND.1PL like that
    'This last sentence, we would say this way'

                  (Escandell-Vidal, 2009: 36)[5]

Also in the Italian speaking children's ClLDs, left peripheral lexical object topics were preferably a-marked and overwhelmingly so (in 88% of their ClLDs: Belletti 2017b; Belletti & Manetti 2018). Some were not, though. They were expressed as a simple DP, with no introducing preposition much as in the adult target language. One example is given in (5) as an illustration (with a null subject *pro* in the clause following the a-Topic):

(5) Il re   lo   sta pettinando
    the king *pro* him-CL is   combing
    'The king, he is combing him'
            (Alice, 4 – *Picture described: Kid combing king* - O *pro* CL V)

Hence, Italian speaking children's productions illustrate a favored, though not obligatory, process of a-marking of direct object left peripheral topics which parallels the distribution of a-Topics in the described variety of Balearic Catalan. Modern Spanish is one step further, with a-marking always obligatory in ClLD, as has been clarified in Leonetti's description mentioned above.

2.3   a-Topics in a diachronic perspective

Indeed, looking at the development of Spanish may be very instructive and inspiring. Following von Heusinger (2008, also relying on Laca's 2006 analysis of the Cid corpus) the picture very clearly emerges that no DOM was present in old medieval Spanish. In contrast, DOM is systematic in modern Spanish as is also clearly revealed by the translation of the same stretch of text. Translation is often an effective magnifier: The following examples taken from von Heusinger (2008) offer a straightforward minimal comparison:

---

5. The example in (4b) is especially interesting in that it shows that the left dislocated object does not need to be animate in order to be a-marked. This suggests that a-marking is a feature in the left periphery that can be associated with topics independently of their animacy, a property often associated to classical DOM.

(6) Old Spanish:
  a. En braços tenedes <u>mis fijas</u>   tan blancas commo el   sol
      in arms   have-2PL my daughters as white   as   the sun
      'In your arms you hold my daughters, as white as the sun'   (Cid, 2333)
  b. Escarniremos <u>las fijas</u>   <u>del</u> <u>Campeador</u>   (Cid, 2551)
      will-humiliate the daughters of the Battler
      'We shall humiliate the Battler's daughters' (von Heusinger 2008: 18a, b)

(7) Modern Spanish:
  a. tenéis   <u>a</u>   <u>mis</u> <u>hijas</u>,   tan blancas come el   sol, en vuestros
      have-2.PL DOM my daughter as white   as   the sun in your
      brazos
      arms
        (*Cantar de mio Cid,* Translation A.Reyes. Madrid: Espasa Calpe 1976)
      'In your arms you hold my daughters, as white as the sun'
  b. y   podremos escarnecer <u>a</u>   <u>la</u> <u>hijas</u>   <u>del</u> <u>Campeador</u>
      and will-can1PL humiliate DOM the daughters of the Battler
        (*Cantar de mio Cid,* Translation A.Reyes. Madrid: Espasa Calpe 1976)
      'We shall humiliate the Battler's daughters' (von Heusinger 2008: 19a, b)

In the same text, however, hence in the same stage of the history of Spanish, von Heusinger (2008, reporting results from Laca 2006: 455) points out that in Cid 80% of (animate) definite direct objects were not not a-marked, whereas 73% of (animate) definite direct objects were a-marked <u>and</u> preposed and doubled by a clitic (in ClLD), as in the following example:

(8) <u>A</u>   <u>las</u> <u>sus</u> <u>fijas</u>   en braço las   prendía   (Cid, 275)
     DOM the his daughters in arm   them hold-3.SG
     'He gathered his daughters in his arms'   (von Heusinger 2008: 19'b)

In other words, a-marking of left dislocated lexical direct object topics is widespread in Cid independently of DOM, which was not available in Spanish at the time. This is a picture that converges in a striking way with the behavior of the Italian speaking children illustrated in (1) and (2) .

  A final concluding remark concerns the further fact presented in von Heusinger (2008) according to which direct object strong pronouns were systematically a-marked in Cid, instead. This is the same situation holding in the non-DOM contemporary Balearic Catalan according to Escandell-Vidal's (2009) description. As mentioned in the introduction, in present day Italian strong personal pronouns allow/require (in the 1st and 2nd person in particular, Beretta 1989; Renzi 1988) a-marking when they are left dislocated, as illustrated in example (9) and similar ones (Beretta 1989; Belletti 2017a):

(9)   A te/*?te   ti   licenziano di sicuro
      to you/you they you-CL fire   for sure
      'They will fire you for sure'

Hence, object strong pronouns tend to be more readily accessible to a-marking. In Italian this only happens in the periphery (with the 1st and 2nd person in particular) and not in the clause internal object position, in the other languages mentioned personal strong pronouns are the first to undergo DOM.[6]

2.4   The development of a-marking

Building on the proposal presented in Belletti (2017a), a diachronic developmental path for a-marking may be identified such that the first instance crucially involves the Topic head in the Left Periphery; a-marking can then affect the lower Topic head in the vP-periphery of the clause and then become a fully grammaticalized marker, expressed in the functional zone of the vP as a *v* component of (accusative) Case marking. This path can be expressed in cartographic terms along the following lines in (10). In (10) the Topic head is endowed with the marker 'a', expressing some interpretative feature that adds up to the Topic/given one. In Belletti (2017a) the proposal is sketched out that left peripheral a-marking adds some affected/involved type interpretation to the topic. For the space reasons I will not further elaborate on the exact characterization of the interpretive value added by 'a' (see reference quoted). Let us simply refer to it as the feature [+ a]. Such feature attracts the object noun phrase into its Spec where the relevant interpretation is implemented in criterial terms (Rizzi 1997). As illustrated in (10a), (10b), the head 'a' then moves to a higher head position thus re-establishing the prepositional order of Italian.

According to this proposal, initially there is a crucial link between a-marking and topicality. Since Topics are characteristically expressed in the left periphery, a-marking is a left peripheral phenomenon in its core manifestation. This is indeed the case for child Italian, Old Spanish, Balearic Catalan.[7] This is the case illustrated by (10a) . When the low topic is involved as in (10b), a-marking is present

---

6. According to Rohlfs (1971), reported in Iemmolo (2010), in the Brussels variety of French and in Languedoc French a-marking is also present and typically affects left dislocated pronouns and only sporadically left dislocated lexical noun phrases in ClLD structures. According to the description in Rohlfs (1971), the Valais-Geneva variety of French is a variety most closely resembling standard Italian in that a-marking is limited to left dislocated personal pronouns only in ClLD structures.

7. Similarly for Sicilian as described in Iemmolo (2007). On the low vP periphery of the clause see Belletti (2004) and references cited there.

(10) a.   ForceP

b.   TP

c.

```
              vP
            /    \
     DP/EA       /\
            v   /  \
             \ /    \
          a   V    DP/IA
```
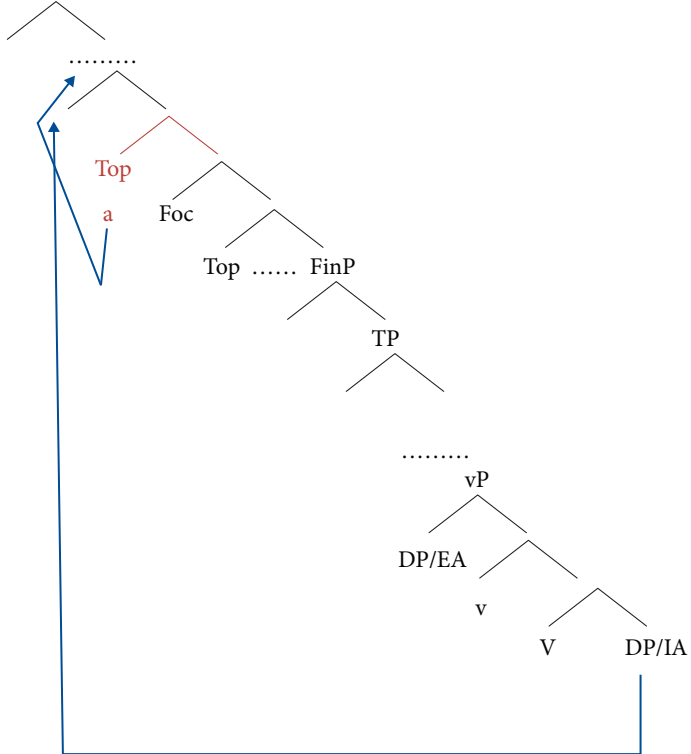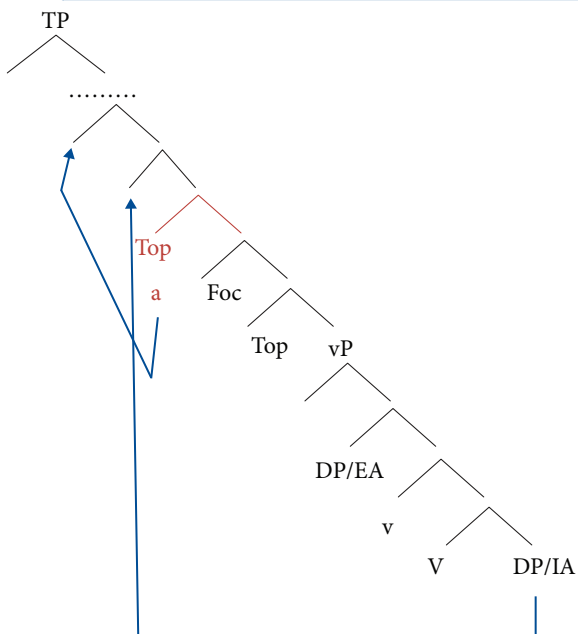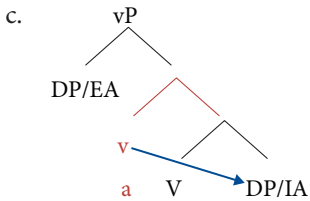
also on non-left peripheral preposed objects, hence it is displayed in the SV(a)O word order as in DOM. Still, some residual topic interpretation is maintained as a-marking of direct objects/DOM is often, characteristically described as marking of a specific/definite object noun phrase. Finally the presence of a-marking is wider and generalized in the fully grammaticalized third stage corresponding to (10c). Following Leonetti's description assumed here, Modern Spanish seems to be in between the latter two steps. It cannot be excluded – nor can it proved though for obvious reasons – that the wide resort to a-Topic emerged in the Italian children's productions could be a sign of a possible language change on the way, eventually leading to some form of DOM in Italian as well.[8] Be it as it may in this respect, the path explicitly describes both the relation with topicality correlating with a-marked preposed objects and its privileged left peripheral realization.

2.5   When objects are Topics

In a language like Italian in the SVO order direct objects are typically the focus of new information (e.g. in narrow focus answers on a question bearing on the object) or part of the focus in a whole new focus sentence (e.g. in the answer to a *what is happening* type question). It is tempting to see a-marking as the way to mark the object when it is not a focus, but rather a topic, in combination with the added left peripheral feature [a], a proposal that I develop in some detail in Belletti (2017a; see also Iemmolo 2010 for related ideas).

It is furthermore tempting to relate the robust fact mentioned in the preceding sections according to which personal strong pronouns are among the first nominal elements to undergo a-marking and then DOM to a property of the Romance pronominal system: in Romance Clitic languages strong pronouns are typically used when the pronoun is focalized (Cardinaletti & Starke 1999). In discourse situations in which the strong form of the pronoun should be preferred (e.g. to put more emphasis on it) while keeping the given/topic interpretation of it,

---

8. Lighfoot (1999) on the leading role of developing children in triggering language change, in this case in the domain of morphosyntax.

a-marking could again be the process resorted to mark the strong pronoun when it is associated with the topic, non-focus, interpretation in the low vP-periphery.
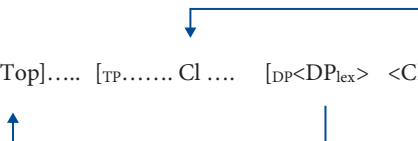
3.   a-Marking on lexical object Topics, not on lexical subject Topics

Italian is a multiple topics language: more than one topic can be present in the left periphery of the clause. For instance a subject DP can be left dislocated together with an object DP:

(11) Gianni Maria la   conosce bene
      Gianni Maria her-CL knows   well
      'Gianni knows Maria well'

The possibility of allowing for multiple topics is acquired early on, as seen in connection with children's productions of sentences like (1) .

In ClLD the sentence following the topic contains a clitic pronouns that refers to it and the clause predicates a property of the topic. Assume for concreteness a doubling type derivation for ClLD whereby a big-DP containing both the lexical DP and the clitic is merged clause internally. The lexical portion of the big DP is moved into the relevant topic position in the left periphery and the stranded remaining clitic part cliticizes into the relevant Clitic head inside the clause (Belletti 2005; Cecchetto 2000; Kayne 2005). The movement + stranding analysis alluded to is schematically represented in (12)

(12)

$$[\dots \dots \quad [\ [\text{Top}] \dots \ [_{TP} \dots \ Cl \ \dots \quad [_{DP} <DP_{lex}> \ <Cl>]]]]$$

When the DP attracted to Topic is the subject of the clause, the stranded pronominal portion, i.e., the resumptive pronoun, corresponds to a silent *pro*, the analogue of an object clitic in a null subject language like Italian (Cardinaletti 2004). (13a) illustrates a derivation in which the DP_{lex} portion of the big DP is moved from a low position of the clause (Rizzi 1982 and 3.1 below), e.g. the position in which it is merged as the external argument of the verb; the stranded portion containing *pro* is moved into the subject position.[9] In (13b) the other possible derivation is illustrated in which the DP_{lex} portion of the big DP is extracted from the big DP in subject position and the portion containing *pro* is stranded in this position.

---

9. To simplify the representation in (13a) *pro* is moved into the subject position. In fact, the moved constituent could correspond to the remnant big DP.

(13)

a [……….[[Top]…..[$_{TP}$ *pro* ….. [$_{DP}$<DP$_{lex}$> <*pro*>]]]]

b [… [[Top]…..[$_{TP}$ [$_{DP}$<DP$_{lex}$> *pro*] ………….]]]

The following considerations apply to both conceivable derivations in (13). Following Rizzi's (2006) and Rizzi & Shlonsky's (2007) approach, a Subject Criterion expresses the interpretive content of an EPP feature in the (relevant, Cardinaletti 2004) subject position in the high portion of the TP.[10] According to the criterial approach this amounts to saying that the clause is about the nominal element filling the subject position.

Topics and subjects have thus much in common: in both cases the sentence following the topic DP and the sentence of which a DP is the subject predicates some property of such DP: the sentence is 'about' the DP in both cases. However, subjects and topics differ in one property: only topics are also given in the previous discourse. In other words, a subject DP is not necessarily also a topic (Rizzi 2005 for discussion). Thus, in cases like (13) in which the subject is also a topic in the left periphery the sentence ends up being about the Topic and the resumptive silent *pro*. It is about a subject that is also a Topic.

Clearly this interpretive option must be available and formally allowed by the grammar as it corresponds to all those cases in which the subject is also a topic. The aboutness relation(s) established with the DP in the left peripheral Spec/TopP and the following clause and the one holding between the *pro*-subject and the (same) clause must count as uniform in the relevant interpretive sense. Since under the criterial approach, the Topic/aboutness interpretation of the topicalized DP is expressed through a Topic-criterion and the aboutness interpretation of the related resumptive *pro* is expressed through the Subject-criterion, we can conclude that the combination of the two criteria on the same referential argument is interpretively congruent in the relevant sense. I would like to entertain the proposal that should a topic DP corresponding to a subject be an a-marked topic this congruence would be lost. The reason is that the interpretive feature that 'a' adds to the topic/given interpretation of the DP in Spec/TopP would not also be present on *pro*. The resumptive *pro* would only express the aboutness interpretation always

---

10. The EPP feature in clauses (Chomsky 2000) is in fact the formal counterpart of the substantive property that all clauses have a subject (Chomsky 1981).

associated with the subject position through the Subject criterion. This would yield a not completely congruent interpretation as the sentence would be about an argument – Topic and subject – that is, in a sense, both 'a' and 'non a'.

No similar problem arises for direct object a-Topics as there is no congruence to be met in this case since there is no Object criterion corresponding to the Subject criterion. Thus, the resumptive object clitic does not satisfy any independent criterion, and the related congruence need with the features expressed by the left dislocated DP in Spec/TopP does not arise. In these cases, the 'a' feature is normally expressed on the only criterial head of the relation, the Topic head, and on its specifier, through the Topic criterion.

3.1   Instances of a-marked subjects

Instances of a-marked subjects are only possible in Italian in the form of experiencer quirky subjects (Belletti & Rizzi 1988; Belletti 2017a).[11] As such, they fill the clause internal subject position and realize a dative, as is also the case in various other languages (e.g. Icelandic, Sigurdsson 2002). Quirky subjects satisfy the Subject criterion exactly as non-quirky ones. A quirky dative subject can also be left dislocated with a resumptive dative clitic present in the following clause. (14a), (14b) exemplify the non left dislocated version and the left dislocated one of the same dative quirky subject (experiencer) in Italian:

(14) a. A Gianni piace la   musica
          to Gianni likes the music
          'Gianni likes music'
      b. A Gianni gli   piace la   musica
          to Gianni to him-CL likes the music
          'Gianni likes music'

The Topic criterion and the Subject criterion are satisfied in both such cases. in (14a) the Subject criterion is satisfied by the dative 'a Gianni'; in (14b) the dative 'a Gianni' satisfies the Topic criterion. As for the Subject criterion two analyses are possible for (14b) depending on the position from which the dative a-DP is moved. I sketch out the two possible analyses, leaving open for the time being the choice between the two as it would require more elaborate discussion: If the dative 'a Gianni' is moved from the high subject position, then according to one analysis the Subject criterion my be assumed to be satisfied by the stranded dative

---

11. The external argument of Italian/Romance causatives can also be a-marked and result in a dative. 'a' is here part of the functional verbal spine realizing a small 'v' that, according to Belletti (2017c) recent discussion is selected by the causative head.

clitic.[12] This analysis is the same as the one illustrated in (13b), modulo the dative nature of the dislocated noun phrase and of its resumptive pronominal. If the dative 'a Gianni' instead is moved from the lower position in the vP domain where it is first merged, then the remnant DP would move into the subject position and the Subject criterion would similarly be satisfied by the stranded dative clitic (see footnotes 9 and 12).[13]

Manzini and Savoia (2005) point out a further instance of a-marked subjects also reported in Rohlfs (1969) occurring in cases in which the subject is long extracted as in the example quoted in (15), taken from the list presented in Manzini & Savoia (2005) as an illustration:

(15)   (ma) ki   t'   kre:d ke 'venga
       [(a)   chi   ti   credi che venga __ (It)]
       (to)   whom you think that comes'
       'Who do think will come?'   (S.Agata Feltria)

To the extent that this type of cases only involves long extraction as in all the quoted examples in Manzini & Savoia's list a natural account interprets these instances of a-marking as the manifestation of (grammaticalized) DOM holding in an ECM-type derivation (along the lines of Chomsky 1981). Manzini and Savoia (2005) propose a somewhat different account, with one main insight that is retained in the analysis just sketched out: despite the fact that a-marking typically affects objects, under some particular circumstances it may, however, also affect a subject. We have just sketched out one such situation. Interestingly, in this situation the a-marked DP is not also interpretively an a-Topic, but rather a DOM-marked DP.

One other instance of an a-marked subject I am aware of is reported in Iemmolo (2010), quoting work on Campidanese Sardinian by Putzu (2008). The example quoted is the following:

(16)   A   chi arriidi   urtimu, arriidi   mellus
       TOP who laugh:PRS.3SG last   laugh:PRS.3SG best
       'He who laughs the last, laughs best'   (Putzu 2008: 412)

---

12. Which should then cliticize on the relevant Clitic head in the same way as a subject clitic (Manzini & Savoia 2005; Poletto 2000).

13. The latter derivation would most closely resemble cases of subject extraction in a null subject language like Italian following the classical analysis of Rizzi (1982), Rizzi & Shlonsky (2007). In these cases a silent expletive *pro* is assumed to satisfy the Subject criterion. According to Rizzi and Shlonsky (2007) subjects can never be moved from the criterial subject position for principled reasons precisely related to the criterial nature of the Subj head. Following the classical analysis of Rizzi (1982), in null subject languages like Italian movement takes place from the low merge position and an expletive null *pro* satisfies the criterion.

Of course, nothing much can be concluded on the basis of one single example that may be the residual manifestation of a *figée* proverb-like expression. However, one property of (16) is potentially revealing: The fact that the Topic subject is here the (free) relative pronoun *chi*. One may then tentatively propose that the gap linked to the a-wh Topic is low in the clause structure, in the vP-merge position; in this position no Subject criterion needs to be satisfied. Hence no interpretive clash arises in this case (an expletive *pro* in the subject should satisfy the Subject criterion, along the lines referred to in footnote 13).

4.   Conclusion

After reviewing the main properties of a-Topics in a non-DOM language like (non southern-central) Italian, the interesting parallel has been highlighted between the linguistic behavior of young Italian speaking children emerged in recent experimental work and comparative data from other Romance languages such as Balearic Catalan and (previous stages of) Spanish. A possible path from a-Topics to DOM has been drawn, building on previous proposals in Belletti (2017a). a-Marking of topics affects direct object DPs. A proposal has been outlined as to why subject DPs could not also be so a-marked when they are topics. According to the proposal subjects can be topics but cannot be a-topics for interpretive reasons ultimately due to the need of simultaneous satisfaction of the Topic criterion and the Subject criterion by the same DP argument. The conclusion is that a subject can be a topic, as made visible in multiple topic languages like Italian, but it cannot be an a-Topic in the general case. Some speculative remarks have been offered in conclusion on some apparent/residual instances of a-marking on subjects, beside those reducible to instances of (dative) quirky subjects.

Acknowledgments

References

Belletti, Adriana. 2004. Aspects of the low IP area. In *The Structure of CP and IP*, Luigi Rizzi (ed.) 16–51. Oxford: OUP.

Belletti, Adriana. 2005. Extended doubling and the vP periphery. *Probus* 17(1): 1–35. https://doi.org/10.1515/prbs.2005.17.1.1

Belletti, Adriana. 2017a. On a-Marking of Object Topics in the Italian Left Periphery. To appear in *From Sounds to Structures. Beyond the Veil of Maya*, Roberto Petrosino, Pietro Carlo Cerrone & Harry van der Hulst (eds), Berlin: Mouton de Gruyter.

Belletti, Adriana. 2017b. Internal grammar and children's grammatical creativity. *Frontiers in Psychology.* doi: https://doi.org/10.3389/fpsyg.2017.02074.

Belletti, Adriana. 2017c. Labeling (Romance) causatives. In *Elements of Comparative Syntax – Theory and Description*. Enoch Aboh, Eric Haeberli, Genoveva Puskás & Manuela Schönenberger (eds), 13–46. Boston: De Gruyter Mouton. https://doi.org/10.1515/9781501504037-002

Belletti, Adriana & Rizzi, Luigi. 1988. Psych verbs and Th-Theory. *Natural Language and Linguistic Theory* 6(3): 291–352. https://doi.org/10.1007/BF00133902

Belletti, Adriana & Manetti, Claudia. 2018. *Topics and passives in Italian speaking children and adults*. Forthcoming in *Language Acquisition*.

Benincà, Paola. 1986. Il lato sinistro della frase italiana. *ATI Journal*, 47: 57–85.

Berretta, Monica. 1989. Sulla presenza dell'accusativo preposizionale nell'italiano settentrionale: Note tipologiche. *Vox Romanica* 48: 13–37.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics. Selected Papers from the XVIII Linguistic Symposium on Romance Languages 1988* [Current Issues in Linguistic Theory 69] Dieter Wanner & Douglas A. Kibbee (eds), 143–170. Amsterdam: John Benjamins.

Cardinaletti, Anna. 2004. Toward a cartography of subject positions. In *The Structure of CP and IP. The Cartography of Syntactic Structures*, Vol. 2, Luigi Rizzi (ed.), 115–165. Oxford: OUP.

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

Cecchetto, Carlo. 2000. Doubling structures and reconstructions. *Probus* 12(1): 93–126. https://doi.org/10.1515/prbs.2000.12.1.93

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

Chomsky, Noam. 2000. Minimalist inquiries. In *Step by Step: Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels & Juan Uriagereka (eds). Cambridge MA: The MIT Press.

Escandell-Vidal, Victoria. 2009. Differential object marking and topicality: The case of Balearic Catalan. *Studies in Language* 33(4): 832–885. https://doi.org/10.1075/sl.33.4.02esc

Friedmann, Naama, Belletti, Adriana & Rizzi, Luigi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119: 67–88. https://doi.org/10.1016/j.lingua.2008.09.002

von Heusinger, Klaus. 2008. Verbal semantics and the diachronic development of differential object marking in Spanish. *Probus* 20(1): 1–31. https://doi.org/10.1515/PROBUS.2008.001

Iemmolo, Giorgio. 2007. La marcatura differenziale dell'oggetto in siciliano: Un'analisi contrastiva. In *Actes du XXV Congrès International de Linguistique et de Philologie Romanes, Innsbruck*, 3–7 September 2007, 341–350.

Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language* 34(2): 239–272. https://doi.org/10.1075/sl.34.2.01iem

Kayne, Richard. 2005. Pronouns and their antecedents. In *Movement and Silence*, 105–135. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780195179163.003.0006

Laca, Brenda. 2006. El objeto directo. La marcación proposicional. In *Sintaxis historica del español. Primera parte: La frase verbal*, Vol. 1, Concepcion Company Company (ed.), 423–475. Mexico: Fondo de cultura económica/Universidad Nacional Autónoma de México.

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics*, 3: 75–114. https://doi.org/10.5565/rev/catjl.106

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Oxford: Blackwell.

Manzini, M. Rita & Savoia, Leonardo, M.. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, Vol. 2. Alessandria: Edizioni dell'Orso.

Manzini M. Rita, & Franco, Ludovico. 2016. Goal and DOM datives. *Natural Language and Linguistic Theory* 34: 197–240. https://doi.org/10.1007/s11049-015-9303-y

Poletto, Cecilia. 2000. *The Higher Functional Fields. Evidence from the Northern Italian Dialects*. Oxford: OUP.

Putzu, Ignazio E. 2008. Per uno studio dell'accusativo preposizionale in sardo antico: Emergenze dallo spoglio del Condaghe di San Pietro di Silki. *Diachronica et synchronica: Studi in onore di Anna Giacalone Ramat*, Romano Lazzeroni et al. (eds), 397–428. Pisa: ETS.

Renzi, Lorenzo. 1988. *Grande grammatica Italiana di consultazione*, Vol. 1: *La frase*. Bologna: il Mulino.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris. https://doi.org/10.1515/9783110883718

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge MA: The MIT Press.

Rizzi, Luigi. 1997. The fine structure of the Left Periphery. In *Elements of Grammar*, Liliane Haegeman, (ed.), 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 2004. Locality and Left Periphery. In *Structures and Beyond the Cartography of Syntactic Structures*, Vol 3, Adriana Belletti (ed.) , 223–251. Oxford: OUP.

Rizzi, Luigi. 2005. On some properties of subjects and topics. In *Proceedings of the XXX Incontro Grammatica Generativa*, Laura Brugé, Giuliana Giusti, Nicola Munaro, Walter Schweikert & Giuseppina Turano (eds), 203–224. Venezia: Cafoscarina.

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh Movement: Moving on*, Lisa Cheng & Norbert Corver (eds) 97–134, Cambridge MA: The MIT Press.

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25(1): 197–226.

Rizzi, Luigi & Shlonsky, Ur. 2007. Strategies of subject extraction. In *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, Hans-Marti Gärtner & Uli Sauerland (eds), 115–160. Berlin: Mouton de Gruyter.

Rohlfs, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti*, 3 Vols. Torino: Einaudi.

Rohlfs, Gerhard. 1971. Autour de l'accusatif prépositionnel dans les langues romanes: Concordances et discordances. *Von ROM zur Romània*, Gerhard Rohlfs (ed.), 61–80. Tübingen: Narr.

Sigurdsson, Halldor A. 2002. To be an oblique subject: Russian vs Icelandic. *Natural Languages and Linguistic Theory* 20(4): 691–724. https://doi.org/10.1023/A:1020445016498

Notes on infinitival relatives in Italian

Guglielmo Cinque and Paola Benincà

Ca' Foscari University of Venice / University of Padova

Here certain properties of Italian infinitival relatives will be discussed which to our knowledge have not been addressed in previous work on this topic (Napoli 1976; Cinque 1988: § 1.1.5; Bianchi 1991, 2007). In particular, we will point out a distinction between two '*da* + infinitive' relative constructions and a difference between *da* infinitival relatives and infinitival relatives introduced by relative pronouns; furthermore, we will illustrate the bounded vs unbounded nature of relativization under restructuring, and the properties of restrictive vs non-restrictive infinitival relatives.

Keywords: *da* and *wh*-infinitival relatives, modal force, island constraints, restructuring

1. Introduction

Italian infinitival relative clauses have been the object of a number of studies (see Napoli 1976; Cinque 1988: § 1.1.5; Bianchi 1991, 2007, and others, who will be quoted in the course of the discussion). On the basis of these analyses Italian infinitival relative clauses appear to be characterized by the following properties:

When the relativization is on the direct object, the relative clause is introduced by the invariant preposition *da*, as in (1):[1]

---

1. In colloquial Italian *da* may also be used when an oblique argument is relativized, provided that the oblique argument is locally bound by a resumptive clitic:

(i) a. Ecco un vaso da metterci   i   fiori freschi
     here a   vase to put-INF = LOC DET fresh flowers
     'Here's a vase to put fresh flowers in.'

   b. Ha   trovato un ragazzo da uscirci assieme tutte le   sere
     have-3SG found a   boy   to go.OUT with   every DET evening
     'S/he's met a guy to go out with every night.'

(1) a. Ho   trovato [un libro$_i$ da recensire t$_i$]
        have-1SG found   a   book to review
        'I have found a book to review.'

    b. [Il   libro$_i$ da recensire t$_i$] è lì sul   tavolo.
        the book to review   is there on-DET table
        'The book to review is there on the table.'

A relative pronoun (*cui* or art. + *quale*) can only be used if it is contained in a larger phrase, as in (2):

(2) a. Ho   regalato a Gianni [un libro$_i$ [con cui$_i$ [passare
        have-1SG given   to Gianni a book with which spend-INF
        la   serata [$_{PP}$ t]]].
        DET evening
        'I have given Gianni as a gift a book to spend the night with.'

    b. Ho   presentato a Gianni [una persona$_i$ [ai cui$_i$ amici
        have-1SG introduced to Gianni a   person$_i$ to whose$_i$ friends
        chiedere consiglio [$_{PP}$ t]].
        ask-INF advice
        'I have introduced to Gianni a person to whose friends he can ask for advice'

In this respect, infinitival relatives behave like finite restrictive relatives (compare (3a) with (3)b), and English infinitival relatives, (see (3)c), as they cannot involve a relative pronoun, (as (3)c), unless the DP is embedded in a PP (see (2)):[2]

(3) a. *Il   vestito il quale/cui hai   comprato non ti
         DET suit   which   have-2SG bought   not 2SG.DAT
        sta bene   (Cinque 1978: 35)
        fit well

    b. *Sto   cercando un libro il quale leggere
        stay-1SG looking a   book which read-INF

    c. I'm looking for a book (*which) to read vs. ...on which to work
                                                        (Pesetsky 1998: 350)

---

2. The rationale underlying this pattern remains unclear. For a review of (some of) the accounts that have been proposed see Hasegawa (1998), who adds another piece to the puzzle from Old and Middle English, where the relative pronoun could be deleted stranding a preposition: *a foot on to goo* (a foot on which to go) or *a hous in to drink and ete* (a house in which to drink and eat). For a stylistically more marked construction admitting object art + *quale* pronouns and pied pipings heavier than PP in finite restrictives and infinitival relatives see Cinque 1982: 281f, 1988: § 1.1.4, § 1.1.5.1, respectively).

The complementizer *che*, which in Italian introduces finite relative clauses on the direct object or the subject (see (4)), as well as the complementizer *that* in English (see the translations of (4)a.-b.), are completely impossible in infinitival relatives (see (5)):

(4) a. Il vestito$_i$ che hai   comprato t$_i$ non ti   sta bene
         DET dress   that have-2SG bought   not 2SG.DAT suit well
         'The dress you have bought does not suit you.'

    b. Il vestito$_i$ che t$_i$ era in vetrina   era molto caro
        DET dress   that   was in shop.WINDOW was very   expensive
        'The dress that was in the window was very expensive.'

(5) *Il   vestito$_i$ che comprare t$_i$ era questo.
      DET dress   that buy-INF   was this

Infinitival relatives can be of the restrictive kind:

(6) a. Lo   studente a   cui   affidare l'incarico è
         DET student to   whom entrust-INF DET-task is
         purtroppo   appena uscito.
         unfortunately just   left
         'The student to whom to entrust the task has unfortunately just left.'
    b. Abbiamo scelto   la stoffa da usare per coprire il divano.
        have-1PL chosen the cloth to use   for cover   DET sofa
        'We have chosen the cloth we can use to cover the sofa.'

Perhaps the most natural examples of restrictive infinitival relatives are those with an indefinite antecedent (see (7)), and they seem to belong to the class of kind-defining relative clauses, as they do not identify the referent of the antecedent but express its characteristics (see Benincà 2012; Benincà & Cinque 2014):

(7) a. Ho   trovato un idraulico a cui   affidare   il   lavoro.
         have-1SG found   a   plumber to whom assign-INF DET job
         'I've found a plumber I can assign the job to.'
    b. Ho   trovato una stoffa da usare   per coprire il   divano
        have-1SG found   a   cloth to use-INF for cover   the sofa
        'I've found a cloth I can use to cover the sofa.'
    c. Ho   bisogno di un assistente da assumere con fiducia
        have-1SG need   of an assistant to hire   with confidence
        'I need an assistant I can hire with confidence.'

2.   Nonrestrictive *da* infinitival relatives?

In Cinque (1988: § 1.1.5), on the basis of sentences like (8), it was assumed that ordinary infinitival RCs are only restrictive:

(8) a. *C'era   persino Giorgio, a cui   parlare di questo.
         LOC-be-IPFV-3SG even   Giorgio, to whom talk-INF of this
     b. *Lida, di cui   essere fieri,   è qua.
         Lida, of whom be-INF proud, is here

It would however seem that some *da* infinitival relatives, like those in (9), can have nonrestrictive usages:

(9) a. Questo libro, da non leggere, parla di   Mozart senza   alcuna
         this   book to not read-INF talks about Mozart without any
         cognizione di causa   (Cinque 1988: 455)
         knowledge of cause
         'This book, not to be read, deals with Mozart with no knowledge of the facts.'
     b. Il   suo consiglio, da prendere sicuramente sul   serio,   è che tu
         DET POSS advice   to take-INF surely   on-DET serious is that you
         non ti   muova.
         not 2SG.ACC move-SBJV.3SG
         'Her/his advice, to be taken seriously for sure, is that you should not move.'

It is to be noted that, while ordinary infinitival relatives are ambiguous between a root possibility ('could') and a root deontic ('should') interpretation (see the next section), the interpretation of these nonrestrictive *da* infinitival relatives is necessarily deontic, which makes one think that they are derived through a reduction from a full finite nonrestrictive relative clauses involving the deontic periphrasis *è da* + infinitive:

(10) a. Questo libro, che   è da non leggere, …
          this   book which is to not read-INF
          'This book, which is not to be read…'
      b. Il   suo consiglio, che   è da prendere sicuramente sul   serio, …
          DET POSS advice   which is to take-INF surely   on-DET serious
          'Her/his advice, that is to be taken seriously for sure…'

3.   Two *da* infinitival relatives in Italian

Like English (object and oblique) infinitival relatives (Bhatt 2006; Hackl & Nissenbaum 2012), Italian infinitival relatives, as noted, can either have a 'could' (possibility) interpretation (cf. (11)a.) or a 'should' (deontic necessity) interpretation when they are introduced by strong determiners (like definite determiners in subject position or universal quantifiers) (cf. (11)b-c.) or contain a negation (cf. (11)d). In some cases both interpretations are available (in fact (11)a. admits a 'should' interpretation as well).[3]

(11)   a.   Ho   finalmente trovato un libro da regalare ai miei figli
             have-1SG finally   found   a   book to give-INF to my children
             'I finally found a book to give to my children as a present.'

      b.   Questo è il   libro da regalare a Gianni.
             this   is DET book to give-INF to Gianni
             'This is the book to give Gianni as a present.'

      c.   Hanno   elencato ogni   libro da mettere all'indice.
             have-3PL listed   every book to   put at-DET   index
             'They have listed every book to be blacklisted.'

      d.   Mi   hanno   segnalato   un libro da non regalare ai
             1SG-DAT have-3PL pointed.OUT a   book to not give-INF to-DET
             miei figli.[4]
             my children
             'I have been shown a book not to be given to my children.'

---

3. Cf. Cinque (1988: § 1.1.5.2). Bianchi (1991: 121; 2007, fn.7) in presenting a sentence like (i) says that infinitival relatives introduced by *da* modifying a subject are always interpreted deontically:

(i)   [Un cane da addestrare] ha   morsicato l'istruttore
       A dog to train-INF   have-3SG bitten   DET-instructor
       'A dog to be trained bit the instructor.'

We agree with the judgment for (i), but we find cases similar to (i) to be acceptable with the possibility reading if they have a generic tense (even (i) perhaps can be marginally interpreted as 'a dog of the kind that can be trained'):

(ii) a.   Un libro da (poter) leggere a letto non può   essere troppo pesante
           A book to (can)   read-INF at bed not can-3SG be   too   heavy
           'A book to (be able to) read in bed cannot be too heavy.'

     b.   Cose da (poter)   fare   senza   spendere troppo si   trovano sempre.
           things to (be able to) do-INF without spending too   REFL find-3PL always
           'Things to (be able to) do without spending too much can always be found.'

4. Giurgea and Soare (2010a: 75) note that negation blocks the possibility reading.

Hackl and Nissenbaum (2012: § 1.3.1) argue that under the 'could' interpretations the infinitival relative clause involves Raising or Promotion of the internal Head, while under the 'should' interpretation the Head can be internal or external. In this paper we do not pursue this aspect of the construction, but point out another difference between the two interpretations. The infinitival relatives with a 'could' interpretation are not (necessarily) islands for extraction ((12)a.) while the ones with a 'should' interpretation appear to be islands for extraction ((12)b.):

(12) a. I   miei figli,   ai quali non ho   un libro da (poter)   leggere
         DET my   children to whom not have-1SG a   book to (can-INF) read-INF
         alla   sera   prima che si   addormentino,…
         at-DET evening before che that REFL go.to.sleep-SBJV.3PL
         'My children, who I do not have a book to read to at night before they go to sleep.'

    b. *I   miei figli,   ai quali ci   sono libri   da non regalare,…
         DET my   children to whom there are   books to not give

4. A difference between the *da* + infinitive and the P *cui*/art. + *qual*- infinitive construction

Clitic Left Dislocation is possible with the latter but not with the former. See (13):

(13) a. Cerco   qualcuno a cui   di questo poter   parlare   con calma.
         look.for-1SG someone to whom of this   can-INF speak-INF with calm
         'I am looking for someone I can talk about this calmly.'

    b. *Cerco   qualcuno da a voi   presentare   al più presto/
         look.FOR-1SG someone to to you introduce-INF quickly/

    b′ *Cerco   qualcuno a voi da presentare al più presto.
         look.for-1SG someone to you to introduce   quickly

5. Another difference between the *da* + infinitive and the P *cui*/ art. + *qual*- infinitive construction

At first sight, in Italian the verb of *da* infinitival relatives cannot be passive ((14)a), while it can in P *cui*/art. + *qual*- (see (14)b) as well as in English (as shown in the English translation of (14)a)). However, in other examples of *da* relative clauses (as (14)c) the passive infinitive seems perfectly grammatical:

(14) a. *Gli   hanno   dato un libro da esser   letto   entro domani
         3SG.DAT have-3PL given a   book to be-INF read-PTCP by   tomorrow

    b.   Cercavano   una medicina con cui   essere curati.
        look.FOR-IPFV.3PL a   medicine with which be-INF cured-M.PL
        'They were looking for a drug with which to be cured.'

    c.   Cerco   un libro da poter   esser letto   in due ore.
        look.for-1SG a   book to can-INF be-INF read-PTCP in two hours
        'I am looking for a book I can read in two hours.'

The ungrammaticality of (14)a, which contrasts with the perfect grammaticality of (14)c containing the modal *potere*, suggests that (14)c is in fact a kind-defining relative clause: the relative in (14)c does not identify the referent of the antecedent, but only expresses its characteristics, as we have seen above in other cases of kind-defining relatives. (14)a qualifies instead as a genuine restrictive relative which identifies the referent of the antecedent.

6.   Are there subject infinitival relatives in Italian?

The answer seems to be negative. Cases which in English are often interpreted as subject infinitival relatives (see (15)) are perhaps more accurately analyzed as either purpose control structures (in Italian they are rendered by a modal finite relative clause (see (16)a)) or, in contexts where the noun is modified by ordinals, superlatives, or *solo* 'only', by an infinitive introduced by the preposition *a* (see (16)b):

(15) a.   The man to fix the sink is here.   (Bhatt 2006: 9)
       b.   The first to walk on the moon visited my school yesterday.   (*Ibidem*)

(16) a.   L'uomo   che deve aggiustare il   lavandino è qua.
          DET-man that must fix   the sink   is here
          'The man who must fix the sink is here.'
       b.   Il   primo a camminare sulla   luna è stato   Armstrong.
          DET first to   walk   on-DET moon is be-PCTP Armstrong
          'The first one to walk on the moon was Armstrong.'

In fact, Williams (1980: § 2.3.2) takes English subject 'infinitival relatives' to involve PRO rather than the trace of an empty operator. The fact that object infinitival relatives ((17)a.) but not subject infinitival relatives ((17)b) permit long-distance extraction in English also seems to suggest that the former but not the latter involve A-bar movement (as the trace is case-marked in the former though not in the latter case):

(17) a.   Here's the book$_i$ to try to get John to read t$_i$ (Bhatt 2006: 12)
       b.   *The man$_i$ for us to try t$_i$ to fix the sink is here   (cf. Bhatt 2006: 6)

Maiden and Robustelli (2013: 141) observe another apparent case of subject infinitival relative in English (*He's not a man to abandon his friends*, meaning 'He is not a man of the kind that abandons his friends/who would abandon his friends') and note that in Italian the same construction is introduced by *da* (*Non è un uomo da abbandonare i suoi amici*), but here there is possibly a silent *tale* "such", which can actually be overt (*..tale da..*); in other words, this is another case of kind-defining relative clause.

A third case of an (apparent) subject infinitival relative in English mentioned in Bhatt (2006: 9) (*The book to be read for tomorrow's class is kept on the table*) has no counterpart in Italian (**Il libro da essere letto per la lezione di domani….*); it is possibly based on the modal construction containing the *is to* modal periphrasis in English (Kayne 2016), which can contain a passivized verb (*This is to be read by tomorrow*), as opposed to the corresponding modal periphrasis *è da* in Italian, which cannot (**Questo è da essere letto per domani*). Maiden and Robustelli (2013: § 7.2.7) note that "in this construction the infinitive may not be passivized […]. However, the passive formed with reflexive *si* is possible" (p. 141): *Questo è da leggersi per domani*. Also see Burzio (1986: 77, fn 36), after Belletti (1982): *Sono cose da farsi al più presto* '(they) are things to do SI as soon as possible'.

7. *Da* infinitival relatives and restructuring

As opposed to English, where (non-subject) infinitival relatives are not clause-bounded (can span across two, or more, clauses) (see (18)), Italian *da* infinitival relatives, as noted in Burzio (1986: 346ff), are clause-bounded (compare (19) with (18)).[5]

$$(18)\text{a. He is a person} \left\{ \begin{array}{l} \text{to regret to have admired} \\ \text{to convince Maria to invite} \\ \text{to suggest that Maria invite} \end{array} \right\} \quad \text{(Burzio 1986: 346)}$$

$$(19) \text{ *E' una persona} \left\{ \begin{array}{l} \text{da rimpiangere di aver ammirato} \\ \text{da convincere Maria a invitare} \\ \text{da suggerire che Maria inviti} \end{array} \right\} \quad \text{(Burzio 1986: 346)}$$

---

5. This is also true of French and Romanian (Giurgea and Soare (2010a. 76; 2010b: § 2), though certain examples appear to be acceptable. See Huot (1981: 171), cited in Abeillé, Godard, Miller and Sag (1996, fn.10):

(i) a. Je cherche un projet auquel lui proposer   de participer
        I look for a project in which   to propose to him to take part

    b. Je ne   vois personne à qui lui   conseiller de s'addresser
        I don't see anyone   to whom to advise   to refer

As Burzio (1986: 346f.) also noted, there are, however, systematic exceptions to the clause-boundedness of *da* infinitival relatives. These are provided by restructuring configurations (but such exceptions are only apparent since restructuring configurations are arguably mono-clausal):[6]

(20) a. C'è   solo una cosa da **dover** fare   per domani.
          LOC-be-3SG only one thing to need do-INF for tomorrow
          'There's only one thing we must do by tomorrow.'

      b. Ho   trovato   un libro da **poter** leggere in vacanza.
          have-1SG find-PCTP a   book to can   read-INF on holiday
          'I have found a book I can read during my holidays.'

      c. L'unica   cosa da **saper** fare   era questa.
          DET-only thing to know do-INF is   this
          'The only thing one needs to know how to do is this one.'

      d. Cercavano   un problema da **riuscire a** risolvere subito.
          look.FOR-IPFV a   problem to manage to solve   immediately
          'They were looking for a problem they would be able to solve
          immediately.'

      e. Se c'è   una cosa da **provare a** fare   subito   è questa.
          if LOC-be-3SG one thing to try-INF to do-INF immediately is this
          'If there is one thing one needs to try to do immediately, it is this one.'

      f. Ho   trovato   qualcosa   da **far**vi   fare.
          have-1SG find-PTCP something to make-INF = 2SG.DAT do- INF
          'I have found something for you to do.'

      g. C'è   una sola cosa da **cominciare** a fare.
          LOC-be-3SG one only thing to begin-INF   to do-INF
          'There is only one thing that we must start doing.'

      h. L'unica   cosa da **continuare** a fare   è questa.
          DET-only thing to continue-INF to do-INF is this
          'The only thing that we must keep on doing is this one.'

      i. L'unica   cosa da **andare a** fare   subito   è questa.
          DET-only thing to go-INF   to do-INF immediately is this
          'The only thing that we should go and do immediately is this one.'

      l. l. C'è   un solo libro da **finire**   di leggere   per domani.
          LOC-be-3SG one only book to finish-INF of read-INF by tomorrow
          'There is only one book that must be finished reading by tomorrow.'

      m. L'unica   cosa da non **tornare**   a fare   è questa.
          DET-only thing to not return-INF to do-INF is this
          'The only thing not to do again is this one.'

_____

6. Sentences similar to some of those reported in (20) are also noted in Napoli (1976: 307).

Yet, not all restructuring configurations qualify as exceptions. There appears to be a generalization: only those restructuring predicates which are lower than *potere/dovere* are possible. All higher ones (according to Cinque's 2006 hierarchy) are impossible. See (21):

(21) a. *Cercava   una cosa da non **sembrare** apprezzare.
          look.FOR-IPFV a   thing to not seem-INF appreciate-INF

      b. *Questa è una cosa da **soler**   fare   con calma.
          this   is a   thing to use-INF do- INF with calm

      c. *Un errore   da non **tendere a** fare   è proprio questo.
          a   mistake to not tend-INF to do-INF is precisely this

      d. *L'unico   lavoro da **finire**   **per** accettare   è questo.
          DET-only job   to end-INF for accept-INF is this

      e. *Se trovate   una cosa da **voler/intendere/desiderare di** fare
          if find-2PL a   thing to want/intend/desire-INF   of do-INF
          ditemelo.
          tell-IMP = 1SG.DAT-2SG.ACC

      f. *L'unica   cosa da **smettere di** fare   è proprio questa.
          DET-only thing to stop-INF of do-INF is precisely this

Acknowledgments

References

Abeillé, Anne, Godard, Danièle, Miller, Philip & Sag, Ivan. 1996. French bounded dependencies. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.41.4381&rep=rep1&type=pdf> (6 February 2018).

Belletti, Adriana. 1982. 'Morphological' passive and pro drop: The impersonal construction in Italian. *Journal of Linguistic Research*, 2(4): 1–33.

Benincà, Paola. 2012. Determiners and relative clauses. *Iberia* 4(1): 92–109. <http://revistas.ojs.es/index.php/iberia> (6 February 2018).

Benincà, Paola & Cinque, Guglielmo. 2014. Post-copular relative clauses in the diachrony of Italian. In *Diachrony and Dialects*, Paola Benincà, Adam Ledgeway, Nigel Vincent (eds), 257–278. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198701781.003.0013

Bhatt, Rajesh. 2006. *Covert Modalities in Non-finite Contexts*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197341

Bianchi, Valentina. 1991. Le relative infinitive e altre strutture modali infinitive in italiano. *Quaderni del Laboratorio di Linguistica* 5: 51–69. Pisa: Scuola Normale Superiore.

Bianchi, Valentina. 2007. Wh-infinitives and the licensing of "anaphoric tense". In *Proceedings of the 'XXXII Incontro di Grammatica Generativa'*, M. Cecilia Picchi & Alan Pona (eds), 35–47. Alessandria: Edizioni dell'Orso.

Burzio, Luigi. 1986. *Italian Syntax. A Government-Binding Approach*. Dordrecht: Reidel.

Cinque, Guglielmo. 1978. La sintassi dei pronomi relativi 'cui' e 'quale' nell'italiano moderno. *Rivista di Grammatica Generativa* 3: 31–126.

Cinque, Guglielmo. 1982. On the theory of relative clauses and markedness. *The Linguistic Review* 1: 247–294. Reprinted, with corrections, in Cinque, Guglielmo, *Italian Syntax and Universal Grammar,* 54–103. Cambridge: CUP. https://doi.org/10.1515/tlir.1982.1.3.247

Cinque, Guglielmo. 1988. La frase relativa. In *Grande Grammatica Italiana di Consultazione*, Vol. 1, Lorenzo Renzi (ed.), 443–503. Bologna: il Mulino.

Cinque, Guglielmo. 2006. *Restructuring and Functional Heads*. Oxford: OUP.

Giurgea, Ion & Soare, Elena. 2010a. Modal non-finite relatives in Romance. In *Modality and Mood in Romance: Modal Interpretation, Mood Selection, and Mood Alternation*, Martin G. Becker & Eva-Maria Remberger (eds), 57–80. Berlin: Mouton de Gruyter.

Giurgea, Ion & Soare, Elena. 2010b. Predication and the nature of non-finite relatives in Romance. In *Edges, Heads, and Projections. Interface properties* [Linguistik Aktuell/Linguistics Today 156], Anna Maria Di Sciullo & Virginia Hill (eds) 191–213. Amsterdam: John Benjamins. https://doi.org/10.1075/la.156.12giu

Hackl, Martin & Nissenbaum Jon. 2012. A modal ambiguity in *For*-infinitival relative clauses. *Natural Language Semantics* 20: 59–81. https://doi.org/10.1007/s11050-011-9075-9

Hasegawa, Iroshi. 1998. English infinitival relatives as prepositional phrases. *English Linguistics. Journal of the English Linguistics Society of Japan* 15: 1–27.

Huot, Hélène. 1981. *Constructions infinitives du français: Le subordonnant* De. Geneva: Droz.

Kayne, Richard S. 2016. Comparative syntax and *IS TO**. *Linguistic Analysis* 39(1–2): 35–82.

Maiden, Martin & Robustelli, Cecilia. 2013. *A Reference Grammar of Modern Italian*. London: Routledge.

Napoli, Donna Jo. 1976. Infinitival relatives in Italian. In *Current Studies in Romance Linguistics*, Marta Luján & Frederick Gerald Hensey (eds), 300–329. Washington DC: Georgetown University Press.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In *Is the Best Good Enough? Optimality and Competition in Syntax*, Pilar Barbosa, Daniel Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds), 337–383. Cambridge MA: The MIT Press.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11: 203–238

Negation and negative copulas in Bantu

Gloria Cocchi

University of Urbino

The main aim of this work is to further speculate on the many syntactic similarities, already discussed in previous work, which may be observed in Bantu and Romance languages. In particular, this work analyses the expression of negation in Bantu, a phenomenon which involves different elements and multiple positions. Crucially, Bantu negation is generally encoded in a specialized prefix, which shows up at the left edge of the complex verbal form; however, negation may also interfere either with tense feature – at least in languages, like Swahili, which exhibit morphologically different tense/aspect infixes in affirmative and negative clauses – or with modality, encoded in the final inflection. This recalls the situation observed also in Romance varieties and especially in Northern Italian dialects.

Keywords: negation, Bantu languages, copulas, Northern Italian dialects, tense/aspect infixes

1.   Introduction

As widely discussed in previous articles,[1] Bantu verbal forms are morphologically complex, and consist of several parts (Alexandre 1981: 361). Besides the root, subject prefix and final inflection are (almost) always present, as well as the tense/aspect (T/A) infix, at least in most finite forms. Other morphemes appear instead in a very restricted number of cases; in particular, the so-called pre-prefix emerges almost only in negative verbal forms, though this is not the sole position where negation may be expressed. Indeed, notable differences may be recorded across languages, as well as in different constructions within the same language.

This work proceeds as follows. In Section 2 we will describe the morphological make-up of copulas and their negative counterparts; indeed some negative forms crucially make use of these elements, as we will see in Section 3, which describes

---

1. See in particular Cocchi (2000, 2003, in press).

negation in Bantu inflected forms. Section 4 contains a comparison between nega-
tion in Bantu and in Romance varieties. Finally, Section 5 summarizes the conclu-
sions reached and the questions which remain open.

The data discussed will generally be drawn from Swahili, the most well-
known and widely spoken Bantu language, and Tshiluba, a rather conservative
core Bantu language spoken in Congo, which we have extensively investigated
in previous work.

2.   Copulas and negation[2]

Most Bantu languages, like Tshiluba, allow the formation of nominal/adjectival
predicates without the intervention of the copula:

(1)   mu-ntu   mu-nene
      CL.1-man CL.1-big
      'the big man' or 'the man is big'

The option in (1) is limited to the unmarked case, i.e. the present tense.[3] Such an
assumption is indirectly confirmed by the Swahili data. In this language, though
the copula would generally be overt, it is nonetheless uninflected: the same form,
ni, is used for all persons, as in (2) :

(2)   mimi / m-tu   ni m-kubwa
      I   / CL.1-man be CL.1-big
      'I am big / the man is big'

The lexical entry for *to be* in Swahili dictionaries (cf. Toscano 1992), is *kuwa*.
This verb is normally inflected in the past, future and present progressive tenses,
though it cannot be employed in simple present clauses, like (2). However, *kuwa*
with copular meaning is a relatively recent formation (Perrott 1957). The most an-
cient Bantu root signifying *be* is *li*,[4] a defective root which cannot be inflected for
tense and exhibits no T/A infix; hence it is typically employed in the generic, de-
fault usage, roughly corresponding to the simple present, where *kuwa* is excluded.

---

2. See Cocchi (2003) for a more detailed discussion.

3. For other tenses, suppletive forms are used, originating from either the verb *kuvua* 'to come',
or the verb *kuakadi* 'to become' (De Clercq 1937).

4. The verb *li* is defective and, among other things, carries no infinitival inflection, i.e. no *ku-*
prefix. As for *kuwa*, its original meaning (sometimes still maintained) is 'to become' (Guthrie
1967/71).

Originally, the Swahili simple present of the copula was inflected for agreement: hence, the *–li* root was preceded by the various subject prefixes (*ni-li, u-li,* etc.). This emerges from a comparison with other, more conservative Bantu languages (Guthrie 1967/71; see also Tshiluba later on). In the following stage the root *–li* was dropped, leaving the sole subject prefixes (*ni, u,* etc.) functioning as copulas.[5]

The differentiated prefixes are nowadays used only in marked cases, e.g. direct questions. Otherwise, the 1st person singular prefix *ni* has extended to cover the whole paradigm. Thus, *ni* is not a copula, as it would appear to be in sentences like (2) , but a subject prefix; the copula is implicit and *ni* is a sort of default form, like the Tshiluba phonetically silent copula in (1) above.[6]

Inflected copulas still survive in the more conservative Tshiluba language. In particular, they are used as emphasised forms in the present (3) , as well as auxiliaries in present progressive forms (4):[7]

(3) mu-ana u-di   mu-nene
     CL.1-boy 3.SG-be CL.1-big
     'the boy *is* big'

(4) mu-ana u-di   u-lua
     CL.1-boy 3.SG-be 3.SG-come
     'the boy is coming'

Turning to negative copular clauses, in Tshiluba the defective *–di* does not have a negative form, and the suppletive root *–ena*, regularly inflected, is used in its place; see the following alternation (Willems 1949: 32):[8]

(5) a. u-di   mu-kese
         2.SG-be CL.1-small
         'you are small'
     b. ku-ena   mu-kese
         neg+2.SG-not to be CL.1-small
         'you are not small'

---

5. This is due to the fact that the [l] sound disappears very easily in Swahili (Perrott 1957).

6. Also in Swahili *ni* may be omitted, especially in idioms (Perrott 1957: 74).

7. In Tshiluba the *li* root is written as *di*, as in this language *l + i > di* [ʤi] regularly.

8. The sole meaning of *kuena* is 'not to be'. Cf. fn. 10 below on the form of the negative 2nd person singular prefix.

In Swahili, the uninflected copula *ni*, used in the unmarked present tense, turns to *si* in negative clauses; originally limited to the 1st person singular, like *ni*, the form *si* has now spreaded over the whole paradigm:

(6) mimi / m-tu   si   m-kubwa (cf. (2))
    I   / CL.1-man be+NEG CL.1-big
    'I am not big / the man is not big'

Besides, in the following section we will see that the form *si* may show up as a negative prefix also with lexical verbs.

3.   Negative inflected forms

Negative inflected forms are somehow complicate, as negation often surfaces in more than one place; in other words, more elements frequently concur to the expression of negation.[9]

First of all, in the majority of cases, negation is expressed in Bantu by means of a pre-prefix, which precedes the subject prefix: this is *ha-* in Swahili and *ka-* in Tshiluba. Regular phonetic contractions are observed when the pre-prefix precedes a vocalic subject prefix, i.e. in 2nd and 3rd person singular.[10]

Also the 1st person singular prefix form is affected by negation, though in this case there is no clearcut phonetic explanation: for all tenses and aspects, in Swahili the 1st pers. sing. negative prefix is *si-*, rather than *hani-*, while in Tshiluba it is *tshi-*, rather than *kan-*. Crucially, these forms do not carry any morphological trace of either the negative pre-prefixes or the subject prefixes; however, it is certainly not accidental that, in Swahili, this prefix coincides with the negative copula *si*, a form which, as seen above, was at first employed only for the 1st person singular.

Among the many inflected forms that we find in Swahili and Tshiluba verbal conjugations, some are negated simply by prefixing respectively *ha-* and *ka-* (or *si-/tshi-* for the 1st pers. sing.). Examples are given by the future tense in Swahili (7) and the perfective aspect in Tshiluba (8) :[11]

---

9. For Swahili data see Perrott (1957) and Ngonyani (2001); for Tshiluba data see Willems (1949) and Burssens (1946).

10. Swahili 2nd pers. sing. *ha + u > hu-* and 3rd pers. sing. *ha + a > ha-*; Tshiluba 2nd pers. sing. *ka + u > ku-*.

11. In Tshiluba verbs, aspect prevails over tense; the form in (8) has been named *aspect constatatif* by Burssens (1946) and has two different variants originating from two different dialects.

(7) a. tu-ta-soma
         1.PL-T/A-read
         'we will read'
     b. ha-tu-ta-soma
         neg-1.PL-T/A-read
         'we will not read'

(8) a. tu-kwatsh-ile / tu-aka-kwata
         1.PL-take-infl / 1.PL-T/A-take
         'we took'
     b. ka-tu-kwatsh-ile / ka-tu-aka-kwata
         neg-1.PL-take-infl / neg-1.PL-T/A-take
         'we did not take'

However, most forms are not negated in this rather simple way. For instance, all Swahili tensed forms except future present a change in the T/A infix, with specialized morphemes for negative forms; see past in (9) and perfect in (10) :

(9) a. tu-li-soma
         'we read'
     b. ha-tu-ku-soma
         'we did not read'

(10) a. tu-me-soma
         'we have read'
     b. ha-tu-ja-soma
         'we have not read'

This strategy, which is quite common across Bantu (Maho 2007), is never employed in the most conservative Tshiluba language, where negation may be expressed instead through the use of the negative copula.[12] See for instance the form in (11) ,[13] roughly corresponding to our present: in affirmative clauses it may be either simple or progressive (i.e. with copula), but in negative clauses it is always progressive, and the negative pre-prefix precedes the negative form of the copula (cf. Section 2):

(11) a. tu-kwata / tu-di   tu-kwata
         1.PL-take / 1.PL-be 1.PL-take
         'we take/we are taking'

---

12. See Cocchi (2003) for a unitary analysis of Bantu copulas and T/A infixes, which are generally mutually exclusive.

13. Burssens (1946) calls it *aspect duratif*. It is the most unmarked form and does not carry a T/A infix.

b.   ka-tu-ena   tu-kwata
    neg-1.PL-be 1.PL-take
    'we don't take' or 'we are not taking'

Very interesting is also the expression of the negative infinitive in Tshiluba:[14] indeed, in this case, the negative morpheme does not precede the prefix (class 15 nominal prefix *ku-*), but follows it. Notice also that the final vowel has changed as a consequence of negation (regular vowel harmony applies too):

(12) a.   ku-kwata / ku-akula
        CL.15-take / CL.15-eat
    b.   ku-ka-kwatshi / ku-ka-kulu
        CL.15-neg-take / CL.15-neg-eat
        'not to take   / not to eat'

Indeed the final inflection, which encodes modality in Bantu,[15] is also affected by negation. We observe a change in the final vowel also in the negative *intentionnel* aspect, a form roughly corresponding to our subjunctive (*tu-kwate > ha-tu-kwatshi*).

    This latter strategy (change in the final vowel in negative forms) is not unknown to Swahili either, as it is employed in the negative present, a form where the T/A infix disappears and the final vowel changes:

(13) a.   tu-na-soma
        1.PL-T/A-read
        'we read'
    b.   ha-tu-somi
        neg-1.PL-read
        'we do not read'

To sum up, this brief and certainly not exhaustive survey of Bantu negative forms establishes that negation is expressed with at least three different types of morphemes, which appear in different positions:[16]

---

14. The negative infinitive is rarely used and it is found in sentences like the following (from Willems 1949: 87):

i.   ku-tumika kudi kulengele, <u>ku-ka-tumiki</u> kudi kubi 'to obey is good, not to obey is bad'.

15. The final vowel is in fact different from -a in the *intentionnel* aspect (see example in the text), or in nominal forms, corresponding to our participles; cf. Willems (1949).

16. For more data on different Bantu languages see Maho (2007), who also identifies three positions (initial, medial and final) for negation.

1. a specialized pre-prefix, which precedes the subject prefix ('pure' negation);
2. some specialized T/A infixes (or copulas), which follow the subject prefix (aspectual negation);
3. a different final inflection, with respect to the affirmative counterpart (modal negation).

Besides, we must remember the peculiar use of a suppletive form, reminiscent of the negative copula, that merges together negation and subject prefix, and which is employed exclusively for the 1st person singular, in all tenses and aspects.

Finally, it is worthwhile underlining that negation often (but not necessarily always) emerges more than once in the same form. This is particularly evident in Swahili, where, with the exception of future, negation is always marked twice: pre-prefix and T/A infix *or* pre-prefix and final inflection (Ngonyani 2001).

4.   A Bantu-Romance comparison

The comparison between Bantu and Romance languages, with a special emphasis on Northern Italian dialects (NIDs), is particularly promising, in virtue of the many similarities that these varieties exhibit.[17] As concerns negation, in the following subsections we will discuss some interesting correspondences which emerge between Bantu and NIDs.

4.1   Jespersen's cycle

The expression of negation in Bantu involves different morphemes and discontinuous positions, as seen above; this recalls the empirical generalization known as Jespersen's cycle, and which is generally illustrated by the case of French:[18]

(14) stage 1   stage 2   stage 3
      ne   ne… pas   pas

Indeed, in Latin, negation was expressed by means of a preverbal particle, *ne*; this situation is preserved in old French (as well as in Italian and Spanish). Later on, the adverbial-like element *pas* – originally a noun meaning 'step' – was employed, in postverbal position, to reinforce *ne* with motion verbs, and was afterwards extended to all verbs. Finally, in recent colloquial French, negation is expressed by the sole *pas*, and the original negative particle *ne* is falling into disuse.

---

17. See Cocchi (2000, 2003, in press); De Cat & Demuth (2008) among others.

18. Cf. Jespersen (1917), discussed in Devos & van der Auwera (2013: 6).

Within Bantu languages, we also have evidence of the three stages. The most conservative Tshiluba, where negation is generally expressed by the sole pre-prefix, would be in stage 1, like Italian. Swahili, where two elements are generally needed to encode negation, would be in stage 2, like Standard French. Finally, stage 3, represented by colloquial French in Romance, also finds a correspondence in Bantu, e.g. in the Ciwoyo dialect of Kikongo, according to the analysis offered by De Kind et al. (2013).

The three stages are also widely represented within NIDs: indeed, in this Romance area, we may find several examples of varieties which involve only the preverbal negation, only the postverbal one, or need both (Manzini & Savoia 2005: 127–141).[19]

4.2   Initial negation

It is a well-known fact that NIDs are characterized by the presence of subject clitics, namely elements that can be analysed as the morphologically free counterpart of Bantu bound subject prefixes (Cocchi 2000 and related work).

Crucially the preverbal negation, also a clitic, may show up, in different NIDs, before or after the various types of subject clitics. Interestingly, however, in most NIDs undifferentiated clitics precede negation (15a). Moreover, the person feature of the clitic also interferes with the positioning of negation: in particular, in many varieties, Person-clitics are indeed the sole clitics which follow negation (15b):[20]

(15) a. a   nə   ðɔrm   (Càsola, in Manzini & Savoia 2005/3: 288)
         CL.S NEG sleep
         '(I) don't sleep
     b. n   tə   ðɔrm
         NEG CL.2SG sleep
         'you don't sleep'

This picture may recall the situation observed in Bantu languages, where the negative morpheme precedes all inflected subject prefixes but it is preceded by the infinitival prefix (cf. (12b) above); crucially the latter, like the NIDs' undifferentiated

---

19. Of course there may also be intermediate varieties, where, for instance, negation is expressed once or twice in different constructions.

20. Manzini & Savoia (2005, Vol.1: 69 ff.), building also on Poletto (2000), describe the various types of subject clitics, among which the undifferentiated ones (i.e. unspecified in terms of person, gender and number, and glossed as clS), which precede (when present) other clitics, and P-(erson) clitics, namely 1st and 2nd person ones (those denoting speaker/listener), which are ordered last, i.e. most adjacent to the verb.

clitic in (15a), does not carry any person feature. Besides, it does not look accidental that the special behaviour exhibited in Bantu by 1st person singular verbal forms, where subject prefixes merge with negation, is reflected by a parallel special behaviour, in NIDs, of P-clitics (15b), which are often preceded by negation and thus are generally adjacent to the verb stem.

4.3   Medial negation

Romance postverbal negation, emerging in languages at stages 2/3 (cf. (14)), is expressed by *pas* in French and by several similar adverbs in NIDs (16) :[21]

(16) a n   'dørmu 'mia   (Sassello, in Manzini & Savoia 2005/3: 149)
      cls NEG sleep   NEG.ADV
      'I don't sleep'

According to Zanuttini (1997), Cinque (1999), Manzini & Savoia (2005), Poletto (2008) and related work, these elements are assumed to be generated in the aspectual domain, and in particular in the specifiers of the various aspectual heads proposed in Cinque's hierarchy.[22]

Bantu languages at the same stage of Jespersen's cycle, as we saw, do not make use of adverbial elements to reinforce negation, but rather employ specialized T/A infixes (cf. (9)–(10) above). As their label says, these are aspectual elements too; the main difference with respect to Romance adverbs is that they are infixes, thus heads, rather than specifiers. However, in line with Manzini & Savoia's (2005) framework, where it is assumed that clitics do not move but are rather generated where we see them, we can propose that Bantu negative T/A infixes lexicalize the aspectual heads in question (cf. also Ngonyani 2001), rather than be generated in their specifier, like their NIDs' counterparts.

Another interesting correspondence is that, in NIDs, negative adverbs generally follow the auxiliary but precede the past participle in compound tenses, as in (17):

(17) a n ø   'mia   dyr'mi   (*ibidem*)
      cls NEG have NEG.ADV slept
      'I have not slept'

---

21. Like French *pas*, they usually derive from nouns indicating small quantities, like Italian *mica*; however, the latter element is optional in Standard Italian, unlike in French and in many NIDs. See Manzini & Savoia (1998, 2005); Poletto (2008).

22. The various adverbs may be found in different specifiers within the aspectual domain, according to the different information they convey: for instance, the adverb *mica* is crucially a 'presupposition trigger' (see Cinque 1976; Pescarini 2009), unlike French *pas* and other NIDs' adverbs, and obviously unlike Bantu negative T/A infixes.

Similarly, Swahili negative tense/aspect infixes, or Tshiluba negative copulas (cf. (11b) above), which provide the same information as the string auxiliary + negative adverb in NIDs, also precede the lexical part of the verb, the stem.

4.4 Final negation

As said above, in Bantu there is a third position where negation may emerge, namely the final inflection. In NIDs, though there is sometimes more than one negative adverb in the same form, there seems to be nothing directly comparable.

However, if we give a closer look at the variegated situation offered by NIDs, we see that there are varieties, in particular in Lombardy, where the postverbal negation takes the form *nɔ*, which resembles the Italian tonic negation *no* rather than a noun originally indicating a small quantity; therefore it probably represents something different. This intuition is reinforced by the fact that *nɔ* may even surface at the end of the complex verbal form, hence not only after the inflected verb/auxiliary but also after a past participle:

(18) l ɔ 'vistu nɔ (Casorezzo, in Manzini & Savoia 2005/3: 139)
him I-have seen NEG
'I haven't seen him'

Therefore, also in NIDs we observe cases where negation surfaces at the very end of the verbal form. According to Cinque (1999), these low adverbs should be in the specifier of a lower functional head, Voice.[23] In this we may find a correspondence with Bantu final negation, which we also assume to lexicalize a lower head where modality is encoded.

5. Conclusions

In this work we have compared Bantu and Romance languages on the expression of negation. In particular, we have concentrated on some aspects which, in our opinion, promisingly suggest a unitary analysis, in the spirit of a Bantu-Romance connection.

Of course, most work remains to be done. On the one hand, negation is a very complex phenomenon, which interferes with many other syntactic aspects. On the other hand, Manzini & Savoia's (2005) seminal work has shown us how the

---

23. Moscati (2010) discusses very similar examples in relation of the interpretation and scope-taking of low negation. However, these issues go beyond the purposes of the current paper and may be tackled in future research.

interplay of several micro-parameters may give rise to an incredibly rich and variegated panorama of different varieties, even in a relatively restricted geographic area like Northern Italy. It is practically impossible to think of a similar extensive survey for the Bantu linguistic family, as it covers an extremely vast area in Sub-Saharian Africa and includes an enormously high number of languages and dialects, which are still greatly unexplored, especially under a syntactic point of view. But we certainly wish that more data on different Bantu languages would indicate if our intuitions are on the right track.

References

Alexandre, Pierre. 1981. (Les langues bantoues,) Tableau d'ensemble. In *Les langues dans le monde ancien et moderne*, Jean Perrot (ed.), 353–375. Paris: CNRS.

Burssens, Amaat. 1946. *Manuel du Tshiluba* [Kongo-Overzee Bibliotheek III]. Anvers: De Sikkel.

Cinque, Guglielmo. 1976. Mica. *Annali della Facoltà di Lettere e Filosofia dell'Università di Padova*, 1: 101–112.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads – A Cross-linguistic Perspective*. Oxford: OUP.

Cocchi, Gloria. 2000. Free clitics and bound affixes: Towards a unitary analysis. In *Clitics in Phonology, Morphology, and Syntax* [Linguistik Aktuell/Linguistics Today 36], Birgit Gerlach & Janet Grijzenhout (eds), 85–119. Amsterdam: John Benjamins.

Cocchi, Gloria. 2003. Copulas in Bantu: Hints for verb movement. *Quaderni del Dipartimento di Linguistica dell'Università di Firenze* 13: 1–14.

Cocchi, Gloria. In press. Bantu class prefixes: Towards a cross-categorial account. In *Linguistic Variations: Structure and Interpretation*, Ludovico Franco & Paolo Lorusso (eds). Berlin: De Gruyter Mouton.

De Cat, Cécile & Demuth, Katherine (eds). 2008. *The Bantu-Romance Connection* [Linguistik Aktuell/Linguistics Today 131]. Amsterdam: John Benjamins. https://doi.org/10.1075/la.131

De Clercq, Auguste. 1937. *Dictionnaire Luba (Luba-Français – Français-Luba)*. Leopoldville (Congo Belge): Missions de Scheut.

De Kind, Jasper; Devos, Maud; De Schryver, Gilles-Maurice & Bostoen, Koen. 2013. Negation markers, focus markers and Jespersen cycles in Kikongo (Bantu, H16): A comparative and diachronic corpus-based approach. *Workshop on Information Structure in Bantu*, Berlin, 10–11 December.

Devos, Maud & van der Auwera, Johan. 2013. Jespersen cycles in Bantu: double and triple negation. *Journal of African Languages and Linguistics* 34/2: 205–274.

Guthrie, Malcolm. 1967–1971. *Comparative Bantu*, 4 Vols. Farneborough, Hants: Gregg International Publishers.

Jespersen, Otto. 1917, *Negation in English and Other Languages*. Copenhagen: Høst.

Maho, Jouni Filip. 2007. The linear ordering of TAM/NEG markers in the Bantu languages. *SOAS Working Papers in Linguistics* 15: 213–225.

Manzini, M. Rita & Savoia, Leonardo M. 1998. Negation parameters and their interaction in Italian dialects. *Quaderni di Lavoro ASIS* 2: 39–60.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni Dell'Orso.

Moscati, Vincenzo. 2010. *Negation Raising: Logical Form and Linguistic Variation*. Newcastle upon Tyne: Cambridge Scholars.

Ngonyani, Deo. 2001. The morphosyntax of negation in Kiswahili. *Afrikanistische Arbeitspapiere* 68: 17–33.

Perrott, Daisy Valerie. 1957. *Teach yourself Swahili*, 2nd edn. Sevenoaks, Kent: Hodder and Stoughton.

Pescarini, Diego. 2009. 'Presuppositional' negation, modality, and the {addressee}. *Padua Working Papers in Linguistics* 3: 22–28.

Poletto, Cecilia. 2000. *The Higher Functional Field*. Oxford: OUP.

Poletto, Cecilia. 2008. On negation doubling. *Quaderni di Lavoro ASIt* 8: 57–84.

Toscano, Maddalena. 1992. *Dizionario swahili (Italiano-swahili, swahili-italiano)*. Milano: Vallardi.

Willems, Émile. 1949. *Le Tshiluba du Kasayi pour débutants*, 2nd edn. Hemptinne St. Benoit (Congo).

Zanuttini, Raffaella. 1997. *Negation and Clausal Structure*. Oxford: OUP.

On gender and number

A psycholinguistic review

Paolo Lorusso

Scuola Universitaria Superiore IUSS Pavia

In the theory of agreement developed by Chomsky (2000, 2001) φ features are undifferentiated, they are organized in a bundle of features, despite the intrinsically different information that each of them carries. However, while person is found to have an autonomous status in many psycholinguistic studies, number and gender show contrasting results: some studies show a crucial difference in the processing of number and gender, others, mainly ERP, do not. We will review the different psycholinguistic findings and we will propose that number and gender simply denote different nominal classes. The difference found in some experiments for number and gender maybe linked to when the feature is made available in the comprehension of the sentence (early/late cues).

Keywords: φ features, gender, number

1.   Introduction

Greenberg (1963), in his typological analysis based on around 30 languages, points to the presence of an implicational hierarchy among morpho-syntactic features, as in (1) .

(1)   Feature Hierarchy: Person> Number> Gender

The main characteristics of this hierarchy is the implicational status of it: that is, if the lowest one (gender) is present in a language, then the features on the top are also present in that language. A morpho-syntactic feature is a property of words that the syntax is sensitive to and which may determine the particular shape that a word has. Features seem to be the core elements of languages that relate sound and meaning through the agreement configurations. The motivation for the Feature Hierarchy in (1) comes mainly from two types of linguistic evidence: the frequency of occurrence and co-occurrence of the features in the world's languages and the observation of certain syntactic phenomena.

In the theory of agreement developed by Chomsky (2000, 2001) features are undifferentiated, they have been assumed to be organized in a bundle of features, despite the intrinsically different information that each of them carries, as the typological literature maintain. Person, number and gender are a feature set that during agreement computation is uniformly dealt with by the formal operation Agree (Chomsky 2000, 2001), which ensures the copying of the relevant feature information from the controller (goal) to the target (the verb). By the psycholinguistic point of view, this has a straightforward consequence in language processing: during feature checking, person, number and gender are accessed as a bundle, presumably treated as a unit under this view.

However, a lot of studies on language comprehension have showed that there is a dissociation on the comprehension of the different features: the dissociation is mainly based on differences in the recognition time and the percentage of mistakes. Many of them focused on the difference in the comprehension of number and gender (Vigliocco et al. 1996; De Vincenzi 1999; Carminati 2005; Barber & Carreiras 2005; among others). Other studies have revealed that person has an autonomous status in the speakers' competence (Nevins et al. 2007; Mancini et al. 2011; Zawiszensky et al. 2016). We will not focus on the role of person, we will concentrate in reviewing the studies that investigated the differences between gender and number.

So, the Feature Hierarchy is a language universal which was originally put forward on the basis of purely (cross)-linguistic evidence, not psychological evidence. The psychological support for it is not surprising and has some relevance in the syntactic representation that feed the Agree operation. The next sections are devoted to describe in some details the psycholinguistic findings and their theoretical implications about the feature hierarchy. In Section 2, we overview the different psycholinguistic studies on gender and number that show contrasting results, in Section 3 we will propose that gender and number are alike since they can be accounted for in terms of different nominal classes, in Section 4 we will sketch a possible explanation for the differences found in some studies for gender and number. Section 5 is devoted to the concluding remarks.

2.   On the Feature Strength Hypothesis: Number over gender

Di Domenico and De Vincenzi (1995) and De Vincenzi (1999) using a cross-modal priming technique,[1] studied Italian sentences where an object clitic pronoun

---

[1]. Cross-modal priming of lexical decision is a well-established method for probing the activation of competing interpretations of lexically ambiguous spoken sequences (Zwitserlood 1989; Gow & Gordon 1995).
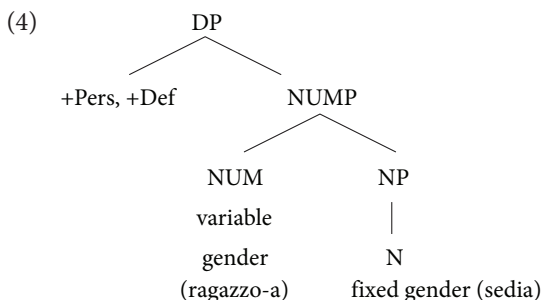
was disambiguated towards one of two antecedents by the gender (2) or the number feature (3) , and obtained significant priming effects only for the latter (3) . Di Domenico and De Vincenzi found a significant priming effect 1000 ms after the offset of the pronoun in the experiment where number (3) was manipulated, but not in the one where gender was manipulated (2) .

(2) il lavoratore   disse alla cuoca   che la padrona di   casa non
     the worker-MASC told the cook-FEM that   the   landlady   not
     poteva sentir-la/-lo
     could   hear -her/ -him
     'the worker told the cook that the landlady could not hear her/ him'

(3) lo sposo disse   agli   alunni   che il
     the   bridegroom-MASC-SING told to the students-MASC-PLUR that the
     vecchio   generale in pensione voleva salutar-lo/-li
     old-MASC-SING general   retired   wanted to greet him/them
     'the bridegroom told to the students that the old retired general wanted to greet him/them'

On the basis of these facts, the authors argue for a modular theory of sentence processing whereby number is processed in an earlier, syntactic stage, and gender at a later stage when semantic information from the lexicon becomes available.

Carminati (2005) in a self-paced reading experiment on Italian indicates that the penalty for disambiguating the antecedent of a null subject is significantly reduced when the disambiguation relies on number compared to when only gender information is available. In particular, she compared the disambiguating power of gender and number and used the superior performance of number over gender as evidence for a 'Feature Strength Hypothesis' according to which there is a correlation between the cognitive significance of a feature and its disambiguating power: the more cognitively important the feature is the better it should be at disambiguating the pronoun that carries it.

But why number is more cognitively salient than gender? Di Domenico and De Vincenzi (1995) assume a serial modular view of processing (Fodor 1983; Frazier 1979) according to which the processor has an input module for syntactic analysis, which operates in the early stages of processing. They reason that because number heads an independent functional projection, it is part of the syntactic representation of the sentence and is therefore 'visible' to the syntactic processor. Their specific proposal is that number is represented autonomously in the lexicon because it carries an independent meaning, while gender is projected in the syntax either with number (variable gender) or with the noun (fixed gender, part of the lexical entry of the noun) as in the syntactic representation in (4) from Carminati (2005: 274).

(4)

```
                        DP
              ┌─────────┴─────────┐
         +Pers, +Def           NUMP
                          ┌──────┴──────┐
                        NUM            NP
                      variable          │
                       gender           N
                    (ragazzo-a)   fixed gender (sedia)
```

For variable and fixed gender Carminati (2005) refers to what psycholinguists define as the semantic and grammatical gender respectively. For semantic gender we refer, in fact, to the inflectional morphemes that refer directly to the sex of the referents.[2] In Italian common gender divisions include masculine and feminine: the gender assignment of nouns can be determined by their meaning biological sex as in (5) . This semantic division is only partially valid, and many nouns may be used to refer to a gender category that contrasts with their morphological gender (5) since in many cases the attribution of grammatical gender is arbitrary (6) .

(5) a. **il**   ragazz-**o**
         the-MAS-SING *ragazz-* MASC-SING
         'the boy'

     b. **la**   ragazz-**a**
         the-FEM-SING *ragazz*-FEM-SING
         'the girl'

(6) a. **la**   sedi**a**
         the FEM-SING *chair*-FEM-SING

     b. *il   sedi-o
         the-MASC-SING *chair*-MASC-SING
         'the chair'

This higher cognitive strength of number over gender is faced by the typological generalization of Greenberg (1966) in the Universal 36, which states: "If the language has the category of gender, it always has the category of number" (Greenberg 1966: 58).

However, results coming from fine–grained methodologies such as electrophysiological studies using event related potentials (ERPs) show that the prominence of number over gender is not that clear, at least in language processing. Event Related Potential (henceforth ERP) has excellent temporal resolution and is therefore ideal for capturing the millisecond-by-millisecond time course of processing.

---

2. For reason of simplicity we will be referring mainly to feminine and masculine gender.

For instance, the significant data that both Di Domenico and De Vincenzi (1995) and Carminati (2005) obtained for number using respectively a priming methodology and a self pace-reading cannot really compete with the array of electrophysiological measures 'related to events'. So, three main ERP component are reported in literature for agreement violations. They are:

– LAN: Left anterior negativity is a negative wave form distributed over the anterior and the left part of the scalp, it occurs between 300 and 400 ms. It is traditionally associated with the processing of ungrammatical information.
– N400: it is a centro-parietal negative wave form that has been considered to respond to semantic, pragmatic or thematic violations.
– P600: is a positive wave form appearing mostly on centro-parietal electrodes. It has been generally interpreted in terms of reanalysis or integration effects in syntactically ungrammatical sentences.

One of the most used technique to investigate the agreement processing has been to compare the response to the processing of grammatical and ungrammatical sentence involving agreement mismatch. The typical pattern found in comparing the ungrammatical with the grammatical situation (ungrammatical sentence minus grammatical sentence) was a pattern of different component: a negative waveform at around 3oo/400 ms (LAN or N400) plus a positive waveform at around 600 ms (P600) after the detection of the agreement violation.

Different studies have investigated the difference between number and gender violations (see Acuña Fariña 2008 for a review): the general pattern found is the one of LAN/N400 + P600 with some variation linked to the type of agreement configuration (Subject-Verb, Noun-adjective, or Determiner-Noun). To our knowledge, the majority of studies that compared number and gender agreement found no differences in their processing across languages (Barber & Carreiras 2005 for Spanish; Osterhout & Mobley 1995 for English; Hagoort & Brown 1999 for Dutch; Gunter et al. 2000 for German; Nevins et al. 2007 for Hindi).

Di Domenico and De Vincenzi (1995) argued for semantic gender as being equivalent to number, so the reason for which many ERP studies have not found any differences between number and gender could be linked to the fact that semantic gender was used in the experimental stimuli. Nonetheless, this is not the case: many studies also compared semantic and grammatical gender and no effect was found (Hagoort & Brown 1999; Gunter et al. 2000; Barber et al. 2004). So, can we really conclude that number has a cognitive power stronger than gender as a disambiguating tool? We need to take a brief excursus on the morphosyntactic and semantic nature of gender and number inflectional class morphology in nouns to partially account for the uniform computation of gender and number.

3. Inflectional noun class morphology

As we have mentioned before, Chomsky argues that in the Agree operation features are undifferentiated, they have been assumed to be organized in a bundle of feature. So the fact that no electrophysiological evidence is found for a differential processing of gender and number features seem to support Chomsky's proposal. So the fact that gender and number are not crucially differentiated in language processing can be linked to how both are interpreted: they represent nominal class morphology that are stored in the speaker's competence.

When we talk about gender specifications in Italian the inflectional pattern is the one reported in Table 1 for the root √ragazz- (boy/girl).

Table 1. Italian gender and number regular inflectional system for the root √ragazz-

√**ragazz-**	**Singular**	**Plural**
Masculine	ragazz-o	ragazz-i
Feminine	ragazz-a	ragazz-e

As first remark, Italian crucially differs from other romance languages such as Spanish where the plural number is lexicalized by –s (chic-o-s, *boys*, chic-a-s, *girls*) and it forms a separate constituent from vocalic endings fixing inflectional class and gender. Plural number has a different inflection, i.e. –*i* and –*e*.

Furthermore, inflectional classes and genders are not isomorphic. As suggested in Franco et al. 2015 (among others), nouns with different inflection –*a* and –*o,* may belong to the same gender, for example the –*o* for feminine (7a) –*a* for masculine (7b). The inflection –*a* can also refer to plurals. The same is true also for plural endings: the –*e* can also be found for both masculine and feminine singular (8a)–(b),–*i* for feminine plural (9) .

(7) a. l-a   man**o**
        the-FEM-SING hand
        'the hand'

    b. il   poet**a**
        the-MASC-SING poet
        'the poet'

    c. le   bracci**a**
        the-FEM-PLUR arms
        'the arms'

(8) a. il   pret**e**
   the-MASC-SING priest
        'the priest'

b. la   mad**re**
the-FEM-SING mother
'the mother'

(9) le   man**i**
the-FEM-SING hands
'the hands'

Since Italian lacks a specialized –*s* morphology for plural, in many cases plurals are not predictable from singulars. Manzini & Savoia (2005) propose that Italian has a dedicated plural morphology –*i,* while other plurals correspond, as in Bantu languages (see Franco et al. 2015; Dechaine et al. 2014) to a switch in nominal class morphology. Bantu languages, in fact, have a total 22 class of nominal classes distinguished for the characteristic feature of the referent (sex, animacy, shape, location etc.). So the syncretism, involving the the endings –*a, –e, –i*, takes place because they represent inflectional class vowels as the nominal class morphology in Bantu: both Italian and Bantu nominal classes are meaningful (Manzini & Savoia 2005, 2007, 2011).

The inflectional morphological endings provide nominal descriptive content to the predicative base of the nominal root (i.e., ragazzo). If gender and number are alike they are both part of the nominal class morphology, they have to be interpreted at LF interface: their interpretation, in fact, in the model of Manzini & Savoia (2005, ff.), triggers the agree operation (not the uninterpretable features on the probe as in Chomsky 2001). The idea is that the predicate represented by the nominal root has an open slot, which has to be filled by its referential properties.

In sum, the so-called semantic gender is the interpreted content of the N inflection but it is part of it as in a nominal class morphology. The representation for *ragazzo/a* boy/girl is like the one in (10) .

(10)

Infl
├── Class
│   ├── √
│   │   ragazz
│   └── Class$_\varphi$
│   [masc]/[fem]
└── Infl
    -*o*/-*a*
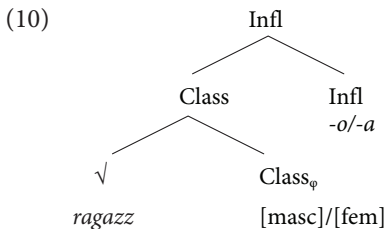
But if there is no difference between gender and number why Di Domenico and De Vincenzi (1995) and Carminati (2005) found a difference in language processing? The first answer is that they use in the stimuli nouns involving the traditional inflectional paradigm in Table 1. Especially for the plurals, if we use the approach of Manzini & Savoia (2010, 2014) for Italian, the exponent -i, connects its semantic

value with its distribution in the Latin nominal paradigms: their proposal is that *-i* (represented as ⊆, the semantic operator indicating inclusion or part/whole) supplies plurality to the noun by isolating a subset of all individuals that are defined by the predicative content of the nominal root (cf. Chierchia 2010). The *–i* morphology influence the general amount of data, since it is the only Italian 'nominal class' that represent uniquely plural referents.

If the nominals used in the stimuli were chosen among the entire range of the Italian nominal class system, it is likely that the effect would have not been that strong. Furthermore, the central role of the definite article inflected for gender and number (found in both experiment), which agree with the inflected nominal may also influence the general results since it implies a sort of reinforced prediction in long sentences like the ones used in the experiments. To our knowledge, no experiment was performed in this respect in Italian.[3] Another factor that seems to interact with the processing of gender and number maybe linked to the position within the sentences of the disambiguating element: whether early or late.

4.   Possible explanations for the different status of number and gender in comprehension tasks

As for the example of the self paced reading experiment of Carminati (2005) which found an higher cognitive strength in disambiguation for number than gender, Acuña Fariña (2008) correctly argues that: Carminati "[…] set the self-paced reading to occur for regions of one clause per bar press: that is, one bar press introduced the whole of the subordinate clause, and the next introduced the whole of the main clause […] and perhaps more importantly, in her materials number information is encountered early, i.e. at the verb, while gender information in appears later, i.e., after the verb" (Acuña Fariña 2008: 406).

In both the experiment, priming (Di Domenico & De Vincenzi 1995) and cue retrieval in a self-paced reading (Carminati 2005) the gender and number cues were used in a configuration antecedent + pronoun. The disambiguation of the pronoun was linked to the difference in the antecedent. The position of the antecedent, at least in Carminati (2005) may have some effect on the results.

---

3. The only exception is the ERP of Molinaro Vespignani & Jobs (2008) in which the authors found an increased P600 in gender mismatch depending on the article selecting the NP where the grammatical DP + NP sentence *lo scialle* (the shawl) is contrasted with two ungrammatical sentences: 1) **il scialle,* where the article is masculine but violate a phonotactic constraint 2) with the feminine article **la scialle* traditional agreement mismatch. The author an incremented P600 for the violation of the phonotactic constraint.

Current psycholinguistic model such as the the cue-based memory model (Lewis & Vasishth 2005) and the self-organized parsing model (Tabor & Hutchins 2004) make different predictions in this respect.

–   The cue-based memory model assumes that comprehenders exploit cues for structure building operations, such as retrieval and reanalysis when these features are realized on the verb, regardless of the processing stage at which these cues are made available.
–   The self-organized parsing model assumes that late cues are less effective than early cues in structure building operations, because the more stable an analysis becomes, the harder to undo (*digging-in effect*).

Predictions are straightforward: the cue-based memory model predicts that late agreement cues will successfully trigger object reanalysis to the extent to which they are realized on the verb, while the self-organized parsing model predicts that late cues will be scarcely effective in triggering reanalysis.

In a preliminary self paced study on ambiguous Italian relative sentences (Villata et al. 2017) we found that when the gender disambiguating cue was late in the sentences it was not effective. For instance, Italian relative clauses are ambiguous between subject and object RCs since Italian permits post-verbal subjects (Arosio et al. 2009, among others). The sentence in (11) is ambiguous between the subject reading in (12) and the object reading in (13) .

(11)   la   bambina che   disegna il   pagliaccio ride
       the girl   that/who draws   the clown   laughs
       'the girl who draws the clown / that the clown draws laughs'

(12)   Subject reading
       $[_{IP}[_{DP}$La bambina$_{iCP}$[che$_{i\ IP}[pro_i[_I$disegna $[_{vP}t_i[_{VP}…[_{DP}$il
       pagliaccio$_j]]]]]]][_I$ride$_i…]]$
       'the girl who draws the clown laughs'

(13)   Object reading
       $[_{IP}[_{DP}$La bambina$_{iCP}$[che$_{i\ IP}[pro_j[_I$disegna $[_{vP}$ il
       pagliaccio$_j[_{VP}\ [_{DP}t_i]]]]]]][_I$ride$_i…]]$
       'the girl that the clown draws laughs'

Here we tested the effectiveness of a late gender cue appearing in a prepositional control sentence after the relative clause in triggering an object RC reanalysis. The sentences are like (14) and (15) .

(14) **Subject reading**

il sindaco$_i$ che consulta la giornalista$_x$ prima di PRO essere
the mayor-MASC that consults the journalist-FEM before of PRO being
ascoltato$_i$ da tutti   vive$_i$ a Parigi
heard-MASC by everyone lives in Paris
'the major who consults the journalist before being heard lives in Paris'

(15) **Object reading**

il sindaco$_i$ che consulta la giornalista$_x$ prima di PRO essere
the mayor-MASC that consults the journalist-FEM before of PRO being
ascoltata$_x$ da tutti   vive$_i$ a Parigi
heard-FEM by everyone lives in Paris
'the major that the journalist consults lives in Paris'

The results showed us that participants were both slower at the spillover region and less accurate in answering the comprehension question in the object condition. Participants failed to access the object analysis 80% of the time.[4]

Early number cue helps in disambiguating between SR and OR (Guasti et al. 2012 and also Villata et al. 2017) while a late gender cue does not, these data in principle support a cue based memory model in which the feature strength hypothesis plays its role (number over gender).

In our discussion about the similarity effects of number and gender in language processing, this excursus on when the cue is presented (the *digging-in effect*) is relevant to account for the effects found in experiments like the ones of Carminati (2005). The fact that only priming and self-paced reading studies present different results for number and gender could be due to how the stimuli are presented and the experiments are designed: the early vs late appearance of the disambiguating cue influences the general results of number over gender (Tabor & Hutchins 2004). Further considerations and data are needed to account for how different nominal classes may influence language comprehension.

5.   Concluding remarks

In this squib we have reviewed the studies that have investigated the role that the different φ features have in comprehension: both in disambiguating and in detecting agreement mismatches. We have not reviewed the data about person, which is almost uncontroversionally found to have an independent status both in syntax

---

4. The detailed results and the statistical analysis are available at <https://www.academia. edu/35784554/Digging-in_effects_in_Italian_relative_clauses>

and in language comprehension. We have been reviewing the results of studies which compared number and gender. On the one side studies like the ones of Di Domenico and De Vincenzi (1995) and Carminati (2005) found differences for number and gender in some comprehension tasks, on the other side ERP experiments have not found any electrophysiological effect confirming the different status of gender and number. We have been arguing that there is no difference in the status of gender and number, but that they are linked to the morphology of different nominal classes (Franco et al. 2015). The differences found with gender and number in comprehension task maybe related to when the disambiguating cue (whether gender or number) was presented in the stimuli (whether early or late in the sentence). To support this observation, we have proposed the results of the preliminary study of Villata et al. (2017) in which the late gender cue rarely triggered reanalysis in contrast to studies where early cues determined full comprehension. However, we need more data to confirm these heterogeneous findings we reviewed in order be able to confirm the proposal that gender an number work alike since they are both part of the nominal class morphology within the speakers' competence.

References

Acuña Fariña, Carlos. 2008. The linguistics and psycholinguistics of agreement: A tutorial overview. *Lingua* 119: 389–424. https://doi.org/10.1016/j.lingua.2008.09.005

Arosio, Flavia, Adani, Flavia & Guasti, Maria Teresa. 2009. Processing grammatical features by Italian children. In *Merging Features: Computation, Interpretation, and Acquisition*, Jose Maria Brucart, Anna Gavarrò & Jaume Sola (eds), 138–155. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199553266.003.0008

Barber, Horacio & Manuel Carreiras. 2005. Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience* 17(1): 137–153.
https://doi.org/10.1162/0898929052880101

Barber, Horacio, Salillas, Elena & Carreiras, Manuel. 2004. Gender or genders agreement? In *On-line Study of Sentence Comprehension; Eye-tracking, ERP and Beyond*, Manuel Carreiras & Charles Clifton (eds), Psychology Press, 309–327, Brighton, UK.

Carminati, Maria. 2005. Processing reflexes of the Feature Hierarchy and implications for linguistic theory. *Lingua* 115: 259–285. https://doi.org/10.1016/j.lingua.2003.10.006

Chierchia, Gennaro. 2010. Mass nouns, vagueness, and semantic variation. *Synthese* 174: 99–149. https://doi.org/10.1007/s11229-009-9686-6

Chomsky, Noam. 2000. Minimalist inquiries. In *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels & Juan Uriagereka (eds), 89–155. Cambridge MA: The MIT Press.

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

Déchaine, Rose-Marie, Girard, Raphaël, Mudzingwa, Raphaël, & Martina Wiltschko. 2014. The internal syntax of Shona class prefixes. *Language Sciences* 43: 18–46. https://doi.org/10.1016/j.langsci.2013.10.008

De Vincenzi, Marica. 1999. Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research* 28: 537–553. https://doi.org/10.1023/A:1023272511427

Di Domenico, Elisa & De Vincenzi, Marica. 1995. Gender and number in the retrieval of pronoun antecedents: Differences in use and representation. In *Actes du deuxième colloque Langues et grammaire*, Lea Nash, George Tsoulas & Anne Zribi-Herts (eds), 95–109, Paris.

Franco, Ludovico, Manzini, M. Rita & Savoia, Leonardo M. 2015. N Class and its interpretation: The neuter in central Italian varieties and its implications. *Isogloss. A Journal on Variation of Romance and Iberian Languages*, 41–68. https://doi.org/10.5565/rev/isogloss.14

Gow, David W. & Gordon, Peter C. 1995. Lexical and prelexical influences on word segmentation. *Journal of Experimental Psychology: Human Perception and Performance* 21(2): 344–359.

Greenberg, Joseph H. 1966. *Universals of language*. Cambridge MA: The MIT Press.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of Human Language*, Joseph H., Greenberg (ed), 73–113. Cambridge, MA: The MIT Press.

Guasti, Maria Teresa, Rotondi, Irene, Vernice, Mirta & Franck, Julie. 2012. The processing of center embedded relative clauses in French and Italian. Poster at AMLaP, Riva del Garda, Italy.

Gunter, Thomas C., Friederici, Angela D. & Schriefers, Herbert. 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience* 12: 556–568. https://doi.org/10.1162/089892900562336

Hagoort, Peter & Brown, Collin M. 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research* 28(6): 715–728. https://doi.org/10.1023/A:1023277213129

Lewis, Richard & Vasishthb, Shravan. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29: 375–419. https://doi.org/10.1207/s15516709cog0000_25

Mancini, Simona, Molinaro, Nicola, Rizzi, Luigi & Carreiras, Manuel. 2011. A person is not a number: Discourse involvement in subject-verb agreement computation. *Brain Research* 1410: 64–76. https://doi.org/10.1016/j.brainres.2011.06.055

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo M. 2007. *A Unification of Morphology and Syntax. Studies in Romance and Albanian Varieties*. London: Routledge.

Manzini, M. Rita & Savoia, Leonardo M. 2010. Case as denotation: Variation in Romance. *Studi Italiani di Linguistica Teorica e Applicata* XXXIX(3): 409–438.

Manzini, M. Rita & Savoia, Leonardo M. 2011. *Grammatical Categories*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511974489

Manzini, M. Rita & Savoia, Leonardo M. 2014. From Latin to Romance: Case loss and preservation in pronominal systems. *Probus* 26(2): 217–248.

Molinaro, Nicola, Vespignani, Francesco, & Job, Remo. 2008. A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research* 1228: 161–176.

Nevins, Andrew, Dillon, Brian, Malhotra, Shiti & Collin Phillips. 2007. The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research* 1164: 81–94. https://doi.org/10.1016/j.brainres.2007.05.058

Osterhout, Lee & Linda A. Mobley. 1995. Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34: 739–773.
https://doi.org/10.1006/jmla.1995.1033

Tabor, Whitney & Hutchins, Sean. 2004. Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30(2): 431–450.

Vigliocco, Gabriella, Hartsuiker, Robert, Jarema, Gonia & Herman H. J. Kolk. 1996. One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes* 11: 407–442. https://doi.org/10.1080/016909696387169

Villata, Sandra, Franco, Ludovico & Lorusso, Paolo. 2017. Digging-in effects in Italian relative clauses, talk presented at Experimental psycholinguistic conference, Minorca 28 June.

Zwitserlood, Pienie. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32: 25–64.

Zawiszewski, Adam, Santesteban, Mikel & Laka, Itziar. 2016. Phi-features reloaded: An ERP study on person and number agreement processing. *Applied Psycholinguistics* 37: 601–626.
https://doi.org/10.1017/S014271641500017X

Micro- and macro-variation

From pronominal allomorphies to the category of
irreality/non-veridicality

M. Rita Manzini

University of Florence

Manzini and Savoia (2005, 2017) argue that morphophonology is involved in
enclisis/proclisis alternations only in so far as it externalizes the syntactico-
semantic category of non-veridicality, as outlined here in Section 1. In Section 2
we review typological literature reporting that the irrealis category governs
the alternation between different pronominal series cross-linguistically. This
evidence potentially fulfills a prediction issuing from the treatment of Romance.
What is more, comparison between treatments of Romance microvariation and
of typological macrovariation reveals a propensity to treat the former in terms of
morphophonological organization and the latter in terms of conceptual systems.
If Manzini and Savoia are correct, efforts at defining opposed notions of macro-
and micro-parametrization are not warranted by the evidence (Section 3).

Keywords: clitics, modality, microvariation, parameters, Romance

1.   Enclisis/proclisis alternations in Romance

In many Romance varieties, proclitic and enclitic position correlate with a number
of morphophonological and order alternations. There are three main such alterna-
tions, namely (a) segmental allomorphies; (b) stress shifts; (c) internal reordering
of the clitic string. In Corsican (1), the vocalic proclitic *u/a/i* alternates with the
syllabic enclitic *lu/la/li*, exemplifying segmental (*l-*) allomorphy.

(1)   a.   **u/a/i**   'cammani
           him/her/them they.call
           'They call him/her/them'
      b.   'camma- **lu/la/li**
           call-   him/her/them
           'Call him/her/them!'   *Zonza*

In Lucanian (2) the same segmental allomorphy as in (1) is exemplified by the alternation of singular *u/a* in proclisis with *lə* in enclisis (final vowels are neutralized in this variety). More importantly, the enclitic shifts the word stress, which in (2b) is on the syllable immediately preceding the clitic, not on the verb root, as in (2a) and also in (1b). In the enclitic group in (2c) the word stress is also shifted to the syllable immediately preceding the *lə* enclitic.

(2)   a.   **u/ a/ lə**   ˈcə:mə
         him/her/them I.CALL
         'I call him/her/them'

      b.   caˈmə- **:lə**
         call- him/her/them
         'Call him/her/them!'

      c.   dana:-**dˈdə-:lə**
         give-to.him- it
         'Give him it!'   *Accettura*

In (3) we return to Corsican to exemplify reordering of the clitic group in enclisis. The order Acc > Dat observed in proclisis in (3a) (the French-type order) is inverted to Dat > Acc in (3b) (the Italian-type order).

(3)   a.   **u/a/i**   **ɖɖi**   ðani
         it.M/IT.F/THEM to.HIM they.GIVE
         'They give it/them to him'

      b.   ˈda- **ɖi-**   **llu/lla/lli**
         give- to.HIM- it.M/IT.F/THEM
         'Give it/them to me/him!'   *Zonza*

Traditional and generative accounts of the phenomena in (1)–(3) are generally conducted within phonological and morphological models. An example of phonological analysis of *l-* segmental alternations is provided by Cardinaletti and Repetti's (2008) treatment of proclisis/enclisis pairs such as *i maɲa* 'they eat' vs. *maɲe-li* 'eat-they?' in the variety of Donceto (Emilia). They argue that *i* is preferred over *li* in proclisis because of the high ranking of *STRUCTURE (imposing economy of structure) over ONSET (imposing the preservation of onset). However in enclisis, *STRUCTURE is ranked lower than *HIATUS (disallowing hiatus) and *VG (Vowel-Glide, disallowing falling diphthongs); the latter constraints are violated if the sequence *maɲa-i* is realized, triggering the form *maɲe-li*. There are open issues in such an account beginning with the fact that an economy restriction such as *STRUCTURE, if construed as strictly phonological, could lead us to have pre-tonic syllables consisting quite generally of the nucleus alone – while we know on typological grounds that this type of syllable is if anything disfavored by

languages of the world. More trivially, we would not expect *li* to appear in proclisis as the dative clitic of Zonza in (4) .

(4) **li**   ð (cf. 3a) ani kwissa
     to.THEM they.GIVE   this
     'They give this to them'   *Zonza*

Phonological treatments are especially popular for the stress shift patterns exemplified in (2) . Bafile's (1993) insight is that a trochaic foot in final position is the preferred metrical option in Romance (except of course for French-type varieties); stress shifts of the type in (2) are therefore viewed as phonological repairs to phonologically ill-formed(/disfavored) metrical structures. The same idea is used by Peperkamp (1996). She argues that if clitics are phonologically inserted at the level of the Prosodic Word, the consequence is that the main word stress rule computes them. In languages without any stress shift, e.g. Italian, clitics adjoin to the Prosodic Phrase, and no reassignment of stress takes place.

But a phonological block against antepenultimate stress is absent in Italian dialects admit, cf. *'fɛmənə* 'women' or *'lavənə* 'they wash' of Accettura. More to the point, antepenultimate stress characterizes instances of kinship noun-possessive clitic combinations, such as (5); (5) is a direct counterexample to a phonological/prosodic treatment of enclitics stress shift, given that the phonological conditions are identical in (2b) and (5), though the syntactico-semantic conditions vary.

(5) 'fratə-   mə/tə
     brother- my/your
     'my/your brother'   *Accettura*

This suggests that these alternations involve the switch between two sets of pronouns, i.e. allomorphs. Even so, Ordoñez and Repetti (2014) take stress to be a key diagnostic for the weak rather than clitic status of the pronouns, in the sense of Cardinaletti and Starke (1999). They propose that in some languages, when the verb is in C, clitics cannot associate with it. In that case weak pronouns attach to the *v*-V complex. This proposal is once again problematic for empirical reasons. There is a considerable amount of syntactic space between the C position taken by the verb in imperatives and the *v* position, that weak pronouns are associated with under Ordoñez and Repetti's proposal. Yet, despite the presence of several intermediate projections no lexical material can appear between the verb in C and the supposed weak pronoun.

Manzini and Savoia (2005, 2017) take a completely different approach. Briefly, vocalic clitics of the form in (1a), (2a), (3a) consists merely of nominal class (gender) and number specifications, which in the appropriate context are sufficient to externalize reference to the 3rd person. In contexts that include the imperatives

in (1b), (2b), (3b) the class and number specifications must be supported by the *l*- lexical base, introducing D properties, as in (6) .

(6) $[_D \, l \, [_\varphi \, u/a]]$

The question is what defines the relevant contexts and why exactly *l*- lexicalization should be matched to them. According to Manzini and Savoia, a prosodic characterization (pre-tonic *vs*. post-tonic position) of the relevant contexts is inadequate, but so is a syntactic characterization based on the position of the verb in I (proclisis) or C (enclisis), following Kayne (1991). The main reason is that in a subset of dialects *l*-allomorphy and reordering is observed when clitics are in the syntactic domain of a negation, as in (7), though proclitic to the verb.

(7)   a.   un **lu/la/li**   cammani
           not him/her/them they.CALL
           'They do not call him/her/them'

      b.   um ɖi   **llu/la/li**   ðani mikka
           not to.HIM it.M/it.F/them give not
           'They don't give it/them to me/him'   *Zonza*

Manzini and Savoia conclude that the definition of the contexts triggering the allomorphies and reordering needs to make direct reference to the interpretive content shared by a modal form of the verb like the imperative and by negation. Irrealis is an obvious candidate. They eventually settle on a category that has independent reflexes in the domain of D lexicalization, namely non-veridicality in the sense of Giannakidou (1998). Specifically, in order for allomorphies and reordering to be triggered the clitics of the relevant varieties need to be in the semantic scope of a non-veridical operator; they further need to be in its syntactic scope. Thus allomorphies and reordering are not observed if the verb remains in I, independently of its semantic import, or if the negation is adverbial (as opposed to clitic) and hence syntactically positioned within vP.

   Vice versa, Manzini and Savoia illustrate in detail how allomorphies are found in all non-veridical contexts characterized by V-in-C including questions, hypothetical and infinitives. We have already mentioned in passing that only in a subset of varieties are the relevant phenomena found with negation; thus they are found in Zonza in (7) but not in Accettura. Manzini and Savoia do not make their account explicit but a very simple, albeit stipulative, option is that anti-veridical (i.e. negative) contexts are not included in the alternations.

   Thus we get to the crucial question why being in the syntactic domain of a non-veridical operator should trigger *l*-allomorphy. We quote: "definiteness is overtly lexicalized by *l*- in those contexts where the surface scope relations, i.e. D lower than the modal verb, do not correspond to the interpreted scope relations,

i.e. definiteness scoping out of modals. The same analysis applies to negation contexts […] lexicalization of the *l-* segment strengthens the visibility of D properties in those contexts where D properties must be read outside the scope of some modal/non-veridical operator that is nevertheless externalized higher than D. We read this requirement as a device to optimize visibility at the SM interface of two key operators (D and C/Neg) interacting at the CI interface" (Manzini & Savoia 2017: 23).

The same analysis holds for stress shift patterns: "the modal V-in-C context requires D properties to be externalized both by an *l-* segment and by prosodic prominence associated with the tonic nucleus immediately preceding *l-*", where the latter is lexically encoded on the *l-* alternant of the clitic. "Stress on the clitic or on the immediately preceding nucleus provides SM salience for the deictic properties that need to be read outside the scope of syntactically higher modal operators" (Manzini & Savoia 2017: 24). We refer the reader directly to Manzini and Savoia's work for the discussion of reordering facts in this perspective.

What matters for present purposes is that "the general picture we obtain is far from the idea that allomorphies and other devices traditionally considered to be morphological simply disrupt the underlying regularity of syntactico-semantic structure. On the contrary they are seen to contribute to the externalization of complex CI information, here the relative scope of modal/non-veridical operators and definiteness" (Manzini & Savoia 2017: 23). In other words, while the surface picture is one of microvariation, involving closely related varieties as well as small segmental, prosodic and order alternations, the explanation requires recourse to the macrocategorial split of (non-)veridicality. Several expectations are borne of such an approach, which the following sections will briefly explore.

2.   Realis/irrealis alternations in the typological debate

In the functionalist/typological tradition of studies, a debate is open on the status of the category irrealis – whether such a category can be consistently defined and how (Mauri & Sansò 2011). This debate opposes exponents of radically skeptical positions on linguistic universals (Cristofaro 2011) to cognitively oriented typologists. What we are interested in is that an important part of empirical evidence comes from alternations in pronominal paradigms triggered by realis and irrealis contexts. If Manzini and Savoia (2005, 2017) are correct, allomorphies, stress shifts and order permutations in Romance clitics are dictated by the same order of categories as the macrovariation described by this typological work – though this potential parallel has not been noticed before by the literature.

We will briefly present just two accounts of the interaction of irrealis and pronominal systems, still routinely quoted in the literature. Roberts (1990) studies the Papuan language Amele. In order to understand the data one has to keep in mind that in Amele "many clauses […] can be linked together which are grammatically dependent in some way on the final clause in the chain. Commonly the verbs in the clauses in the chain [termed medial clauses] will be marked for switch-reference, that is, they indicate if the subject of the following verb is the same as or different from the subject of the marked verb. Also these verbs are normally not marked overtly for categories such as tense or mood but rather these categorial specifications are read off the verb in the final clause and apply to all the verbs in the clause chain" (Roberts 1990: 370). What is relevant here is that "with the simultaneous DS medial verb there are two different sets of subject agreement markers […] which can be selected according to the modal status of the final verb" (Roberts 1990: 371).

In (8) we provide just two examples. The *-en* suffix of the dependent verb in (8a) is what we would call a subject clitic, carrying not just information about phi-features, namely 3rd singular, but also about the fact that it is disjoint in reference from the subject of the superordinate verb (Different Subject, DS). In addition to this, it carries the information that the construct is Realis (R). Compare this with (8b), which is point by point identical to (8a), except that the (8a) is in the past, while (8b) is in the future hence Irrealis (IR). The dependent verb subject enclitic takes the different form *-eb*.

(8) a. Ho bu-busal-en   age qo-in
         pig SIM-run out-3SG.DS.R 3PL hit-3PL.REM.P
         'They killed the pig as it ran out'
     b. Ho bu-busal-eb   age qo-qag-an
         pig SIM-run out-3SG.DS.IR 3PL hit-3PL-FUT
         'They will kill the pig as it runs out'   *Amele*(Roberts 1990: 371)

At this point one may still feel that Romance and Amele are too different to be compared and that the parallel between the Romance data and (8) is at best suggestive. Consider however the list of environments that governs the alternation in (8) : "the verb final categories that co-occur with the realis marked medial verb include all the past tenses […] as well as the present tense. The verb final categories that co-occur with the irrealis marked medial verb include the future tense and imperative, prohibitive, counter-factual, prescriptive, hortative and apprehensive moods" (Roberts 1990: 371). What is more "negation is another category that could possibly have modal status, but in Amele negation is […] outside of the modal system since it […] does not interact with the medial […] realis-irrealis verb" (Roberts 1990: 378). Therefore the contexts that trigger the subject clitic

allomorphy in Amele are eminently compatible with the Romance ones – and the fact that the negation is not a trigger recalls the Romance fact that it is a trigger only in relatively small subset of Romance varieties.

Roberts (1990) goes on to detail the same kind of phenomenon in a number of Papuan languages, beginning with Nobonob. In this language, the realis/irrealis marking is somewhat richer; one series of subject clitics alternates according to the schema in (9) . In other words, morphologically speaking the alternations can take (among others) the form of segmental enrichment of the irrealis form.

(9)		1SG	2SG	3SG	1DU	2/3DU	1PL	2/3PL
	Realis	-i	-e	-e/∅	-ud	-eh	-ut	-eg
	Irrealis	-pi	-pe	-eb	-pu<u>d</u>	-pe<u>d</u>	-put	-peg

Nobonob (Roberts 1990: 379)

In order to establish that all of this is not casual coincidence we need at least a third point of comparison. This is provided by Caddo (a Native American language spoken in Oklahoma) as analyzed by Chafe (1995). The data of Caddo are simpler to process than those of Amele. Quite simply, the Caddo verbs are include a 'pronominal prefix' which "distinguishes person, case, and – of most interest here – reality […]. The reality distinction is an obligatory, clearly marked, and unambiguous feature of every pronominal prefix (with one minor exception), and thus of every verb" (Chafe 1995: 351). There is a single complicating factor namely that all Irrealis contexts (with the exception of yes-no questions) "involve the simultaneous presence of a so-called prepronominal prefix: a verb-initial prefix whose function is compatible with irreality" (Chafe 1995: 353). Unsurprisingly given what we have seen so far, the contexts that trigger the irrealis alternation, in the terms of Chafe, are: Yes-No Questions, Negations, Negative Imperatives, Obligations, Conditions. We provide just one example, involving the negation in (10) .

(10)	kuy-	t'a-	yi=bahw
	NEGATIVE-	1ST.AGENT.IRREALIS-	see
	'I don't see him'		*Caddo* (Chafe 1995: 355)

As the negation again makes its appearance in the Irrealis set, Chafe nevertheless must face the fact that neither futures nor (positive) imperatives figure in the Irrealis set. One of the solutions he proposes is exquisitely syntactic, though expressed in externalist historical terminology: "it is likely that ways of expressing imperatives and futures were established in the language well before the negative and other particles came to be attached to […] verbs. Already entrenched in the language, they would then have failed to participate in the more recent grammaticization of irreality in the pronominal prefixes" (Chafe 1995: 359). Translated in present terminology, one would say that the lexicalization of the irrealis series of

clitics depends not only on the presence of the relevant semantic categories, but also on their syntactic embodiment by the preverbal prefixes (clitics) like the negative one in (10) .

Though variation is found both within the same family (Romance) and across different language families, each single behavior or point of variation illustrated for Romance is replicated by one or the other of the languages considered. It is simply impossible that this kind of distribution is arbitrary – therefore the only question is which explanation has the best empirical coverage both in terms of common features and in terms of possible parameters. Because of obvious limitations of space we cannot delve much deeper into this issue. Nevertheless some general remarks are suggested by the discussion so far. First, The typological debate seems to be polarized between a denial of broad conceptual universals such as (ir)realis (Mauri & Sansò 2011; Cristofaro 2011) and their affirmation (Roberts 1990; Chafe 1995). The generative approach is more articulated since conceptual universals can be seen only through their embedding in computational processes and the externalization at SM interface.

In the light of the typological debate, it is also useful to look back at the debate on Romance. The interesting insight to be gained here is that microvariation studies, concentrating on a great number of closely related varieties (dialects) traditionally consider lowest level (morphophonological) categories best suited to account for the fine grained picture. Typological work, seeing the broad picture spread across an entire languages family (e.g. Austronesian) invokes cognitive categories whose very wide coverage lies open to falsification from more finely detailed data. Eliminating the distortion involves recognizing the macrocategories at play in Romance microvariation. Vice versa, we expect that once the common features are noticed, one may use languages with extremely well-known microvariation to settle the finer points of the analysis cross-linguistically.

3.   Microvariation and macroparameters

There is a sense in which it is useful to speak of microvariation, involving closely related languages and circumscribed areas of the lexicon/grammar – leaving the rest of the lexicon/grammar constant. Microvariation may further be contrasted with macrovariation, including unrelated languages and language families. The final question we briefly raise here is whether it is equally meaningful to talk about microparameters and macroparameters as theoretical constructs – and whether the kind of evidence that we have been reviewing sheds any light on this question. Here is what Kayne (2000: 6) has to say on microvariation and microparameters: "the technique of examining a large number of very closely related languages

promises to provide a broad understanding of parameters at their finest-grained (microparameters), that is, to provide a handle on the question, What are the minimal units of syntactic variation?" It seems therefore that if we follow Kayne, since microparameters are the minimal units of syntactic variation, they are the only real parameters.

Prominent generative scholars have argued against this position arguing not so much against the notion of microparameter as in favour of its coexistence with macroparameters. However an approach like Biberauer et al.'s (2014) does not differ substantially from Kayne's. For Biberauer et al. microparameters and macroparameters represent different levels of application of a given parameter. Thus a given yes/no choice (a parameter) as applied to all/no functional categories defines a macroparameter. When we restrict it to, say, a single functional category it defines a mesoparameter. As we go down the scale, we may further encounter microparametric values which apply to members of a given category. In other words, there is a single unit of variation (a choice open in the lexicon/grammar) and the definition of micro- and macro-parameter is purely extensional – in other words the same parameter may yield micro- or macro- (or meso-) variation.

Finally, Manzini and Savoia (2011), discussing auxiliary selection, have this to say: "It is evident that, to the extent that the primitives manipulated by variation are macrocategories like transitivity or voice, we could describe our approach as macroparametric – though the fact that the unit of variation can be as small as a single lexical item qualifies it as microparametric". In fact, Manzini and Savoia (2011) are more explicit, stating that "macrophenomena can be decomposed into the same elementary conceptual components that determine local lexical variation – and in fact the latter is the true matrix of perceived macroparameters". In other words, macrovariation and macrovariation depend on the same parameters. Kayne (2000) called them microparameters because of their minimal nature. Manzini and Savoia (2011) prefer to insist on the macroparametric nature of their approach in that it involves macrocategorial splits even if applied to a small lexical domain.

In conclusion, Manzini and Savoia's (2011) approach provides a particularly good fit to the empirical findings of this contribution. Though variation involves just morphemic alternations (microvariation), the theoretical splits underlying them involve macrocategories. A corollary of this way of seeing things is that we ought to find that microvariation (among dialects) and macrovariation (within and among families) display the same kind of overall implicational phenomena. This indeed seems to be true of pronouns in the scope of irrealis/non-veridical operators.

References

Bafile, Laura. 1993. Fonologia prosodica e teoria metrica: Accento, cliticizzazione e innalzamento vocalico nel napoletano. PhD dissertation, Università di Firenze.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In *Measuring Grammatical Complexity*, Frederick Newmeyer & Laurel Preston (eds), 103–127. Oxford: OUP.

Cardinaletti, Anna & Repetti, Lori. 2008. The phonology and syntax of subject clitics in Interrogative Sentences. *Linguistic Inquiry* 39: 523–563.
https://doi.org/10.1162/ling.2008.39.4.523

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

Chafe, Wallace. 1995. The realis-irrealis distinction in Caddo, the Northern Iroquoian languages, and English. In *Modality in Grammar and Discourse* [Typological Studies in Language 32], Joan Bybee & Suzanne Fleischman (eds), 349–366. Amsterdam: John Benjamins.
https://doi.org/10.1075/tsl.32.15cha

Cristofaro, Sonia. 2012. Descriptive notions vs. grammatical categories: Unrealized states of affairs and 'irrealis'. *Language Sciences* 34: 131–146.
https://doi.org/10.1016/j.langsci.2011.08.001

Giannakidou, Anastasia. 2011. Negative and positive polarity items. In *Semantics: An International Handbook of Natural Language Meaning*, Vol. 2, Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds), 1660–1712. Berlin: De Gruyter.

Kayne, Richard. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686.

Kayne, Richard. 2000. *Parameters and Universals*. Oxford: OUP.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo M. 2011. *Grammatical Categories*. Cambridge: CUP.
https://doi.org/10.1017/CBO9780511974489

Manzini, M. Rita & Savoia, Leonardo M. 2017. Enclisis/proclis alternations in Romance: Allomorphies and (re)ordering. *Transactions of the Philological Society* 115: 98–136.
https://doi.org/10.1111/1467-968X.12093

Mauri, Caterina & Sansò, Andrea. 2012. What do languages encode when they encode reality status? *Language Sciences* 34: 99–106. https://doi.org/10.1016/j.langsci.2010.11.004

Ordonez, Francisco & Repetti, Lori, 2014. On the morphological restrictions of hosting clitics in Italian and Sardinian dialects. *L'Italia Dialettale* LXXV: 173–198.

Peperkamp, Sharon, 1996. On the prosodic representation of clitics. In *Interfaces in Phonology*, Ursula Kleinhenz (ed.), 102–127. Berlin: Akademie Verlag.

Roberts, John R. 1990. Modality in Amele and other Papuan languages. *Journal of Linguistics* 26: 363–401. https://doi.org/10.1017/S0022226700014717

CHAPTER 8

Concealed pseudo-clefts? Evidence from a Lombard dialect

Diego Pescarini and Giulia Donzelli

CNRS, University of Côte d'Azur / University of Zurich

This paper focuses on the syntax of clefts in the Lombard dialect of Comun Nuovo (Bergamo). In this dialect, clefts are highly constrained (in particular, they are ungrammatical in questions) and, in the contexts where clefts and pseudo-clefts alternate, the distinction between the two is often blurred. We argue that Comunuovese clefts are better analysed as concealed pseudo-clefts (Paul 2001 a.o.).

Keywords: clefts, pseudo-clefts, northern Italian dialects, interrogatives

1. Introduction

The paper investigates the syntax of clefts and pseudo-clefts in the dialect of Comun Nuovo (hereafter, CN).[1] Since in CN clefts are highly constrained, we argue that they are better analysed as (a kind of) pseudo-clefts (in the spirit of Percus 1997; Paul 2001 a.o.).

The structure of the paper is the following: § 2 introduces a simplified taxonomy of clefts and pseudo-clefts and, through a comparison of CN and Italian data, highlights some peculiarities of CN clefts; § 3 deals with interrogation in CN, while § 4 focuses on cleft interrogatives; § 5 elaborates on some theoretical consequences; § 6 concludes.

2. Types of clefts

Clefts are focus constructions in which the copula introduces a focalised XP (on Italian, see Sornicola 1988; Salvi 1991). Following Belletti 2008, we assume that the

---

1. For a corpus-based study on Bergamasco, see Valentini (2012).

XP is extracted from a (small) CP, see (1). Pseudo-clefts, by contrast, are copular constructions in which the subordinate clause is a Relative Clause (RC), either headless or headed. In the former case, we obtain so-called *wh*-(pseudo)-clefts of the type *what he likes more is money* (in Italian, *wh*-pseudo-clefts are rare as headless RCs are seldom introduced by wh elements). In the latter case, see (2), the RC is introduced by a D element or a generic noun, e.g. *the thing that, the person that*, etc., giving rise to so-called *th*-(pseudo)-clefts (Collins 1991).

(1) a. It is *money* [$_{small\ CP}$ that he wants ~~money~~]
     b. è *il denaro* [$_{small\ CP}$ che vuole ~~il denaro~~]

(2) a. Money   is [$_{headed\ RC}$ *the thing* he   wants ~~the thing~~]
     b. Il denaro è [$_{headed\ RC}$ *ciò*   che vuole ~~ciò~~]

In what follows, we focus on the comparison between clefts and *th*-pseudo-clefts. It is worth noting that in many Romance languages the distinction between the two is almost blurred because both may exhibit the same word order, e.g. *copula XP CP/RC*. Despite their similarity, however, clefts and pseudo-clefts differ under several respects. First, in Italian, pseudo-clefts are not necessarily corrective and, in fact, can be used as answers, see (3a). By contrast, clefts often yield a corrective interpretation – save for cases in which the subject is focalised – and, therefore, are slightly degraded as answers to a question, see (3b):

(3)   - chi hai salutato?
      'who did you greet?'
      a. È   Carlo quello che ho   salutato
          It is Carlo the one that I have greeted
      b. #È   Carlo che ho   salutato
          It.is Carlo that I.HAVE greeted
      'It is that book that I gave to Mario'

Second, while clefts allow the extraction of PPs from the embedded clause, (4), *th*-pseudo-clefts cannot have the form *[PP *be* RC], see (4b): [2]

(4) a. È   a Giulia *che* ho   prestato il   mio libro
         it.is to G   that I.HAVE lent   the my book

---

2. Locative PPs can in fact occur in a copular construction along with a RC, e.g.

(i) È   (a) Milano *il posto*   che amo
     it.is (in) Milan   the place that I.LOVE

However, (i) has only a locative interpretation, i.e. 'the place is *in* Milan', not the identificational interpretation characterising pseudo-clefts, i.e. 'the place *is* Milan'.

    b. È   (*a) Giulia *la   persona a cui*   ho   dato il mio libro
        it.is (to) G   the person to whom I.HAVE lent the my book

Surprisingly, CN differs from Italian under both respects. First, clefts are not necessarily corrective. For instance, a cleft such as (5) is a fine answer to a wh question:

(5) Q: Cosa t'   et   dacc al   Mario?
          What you= have given to.the Mario
          'what did you give to Mario'
     A: a l'è chel liber lé   che g   o   dacc[3]
          a=it=is that book there that to.him= I.HAVE given
          'It is that book I gave to Mario'

Second, in CN not only pseudo-clefts, but also clefts cannot have a PP in focus, cf. the following Italian/CN pair:

(6) a. È   da Luca che ho mangiato ieri   It.
      b. *a l'è   da Luca che ho mangiat ier   CN
         a=it=is at Luca's that I.have eaten yesterday

The ban on PP-clefts in CN is quite puzzling under the usual raising analysis of clefts, in which the Focus position is category-neutral.

    The third puzzle about CN clefts regards interrogatives, as it turns out that cleft interrogatives are often ungrammatical, while pseudo-clefts are always grammatical. This restriction is rather surprising as clefting is normally regarded as a common interrogation strategy in northern Italo-Romance (Poletto & Vanelli 1997 a.o.).

(7) a. Chi é-l *(chel)   che (a)l dorma?
         Who is=it the.ONE that he=sleeps
      b. Chi é-l *(chel)   che t-é   vest?
         Who is=it the.ONE that you=have seen

---

3. The behaviour of the particle *a* in Lombard dialects is a well-known puzzle that cannot be addressed here for space limits. For the sake of clarity, in the glosses we try to distinguish the particle *a*, which occurs before subject clitics under certain pragmatic conditions, from the prosthetic vowel *a-* that syllabifies the 3rd person subject clitic. Crucially, the two are in complementary distribution (*a al); our glosses reflect the intuition of speakers.

    The same particle, which arguably derives from a pronominal form (Lat. EGO 'I'?) is attested in several Northern Italian dialects with various functions/interpretations, some of which are related to information structure; for a syntactic analysis of the particle *a* in Paduan, see Benincà 1983/1994.

The following sections illustrate some properties of CN interrogatives and focus on the interplay of clefting and interrogation.

3.   An aside on interrogatives in CN

CN interrogatives are characterised by residual subject clitic inversion. Inversion is forbidden in yes/no questions and, in wh questions, is restricted to present-tense clauses with a deictic temporal interpretation such as (8a). Conversely, enclisis cannot co-occur with present forms with a futurate or habitual interpretation such as (8b) or with past/future tenses, see (8c).

(8) a. ndo   core-l?
         where is.running=he
         'where is he running?'
     b. *ndo   al   cor   a nedal/töcc i martedè sira?
         where he= runs at Christmas/every Tuesday evening
         'where does he run?'
     c. *ndo   al-ha   corìt/corerà?
         where he=has run/will.RUN
         'where did/will he run?'

The second main feature of CN interrogatives is the occurrence of different classes of wh elements, each exhibiting a peculiar syntactic behaviour (see Manzini & Savoia 2011 for similar data on Lombard dialects). For instance, the wh corresponding to 'what' has three possible forms: [sa], [ˈkɔza], and [koˈzɛ]. [sa] belongs to the class of clitic wh elements (in the terms of Poletto 2000; Poletto & Pollock 2006; but see Manzini & Savoia 2011 and Manzini 2014); clitic wh elements in CN can co-occur with subject clitic inversion; [ˈkɔza] never triggers inversion and occurs either *in* or *ex situ*; [koˈzɛ], by contrast, cannot be fronted.

The interplay between clitic inversion, *in/ex situ* placement, and various kinds of wh items gives rise to several patterns of interrogation that, for space limits, cannot be discussed here (see Donzelli 2017).

4.   Cleft interrogatives

This section deals with the distribution of clefts among different types of interrogative clauses. For the sake of clarity, we distinguish three types of interrogatives: *what/who* interrogatives, where a bare DP is interrogated; temporal interrogatives, where a Measure Phrase (MP) is interrogated, and other types of interrogatives.

4.1   who/what interrogatives

CN permits the interrogation of pseudo-clefts, whereas cleft interrogatives are not possible.

(9) a.   Chi é-l   *(chèl) che l è   dre a durmì?   *who*-S
          who is=he   the.ONE that he=is sleeping

     b.   Chi é-l   *(chèl) che ta   set dre ad ardà?   *who*-O
          who is=he   that   that you= are watching

     c.   Cos é-l   *(chèl) che ta   manget?   *what*-O
          what is=it that   that you= eat

It is worth noting that pseudo-clefts do not behave like plain wh interrogatives. First, they do not allow wh in situ and always exhibit inversion, even if the fronted wh element is not clitic, cf. *cos* in (9c). With past or future tenses, where inversion is not permitted, pseudo-clefts are degraded, see (11).

(10)   Chi *l'è/é-l   chèl   che l è   dre a durmì?
       who   he=is/is=he the.ONE that he=is sleeping

(11)   ˀCosa l éra   chèl che ta   séret dre a mangià?
       what it=was that that you= were eating

4.2   Temporal interrogatives

Temporal clefts exhibit a peculiar behaviour as the temporal Measure Phrase can occur as either a DP or a PP (see Benincà 1978 on Italian). In the former case, the copula may agree in number with the MP. In these respects, Italian and CN do not differ. It is worth noting that in CN, as well as in most Northern Italian dialects, number agreement is marked on the subject clitic (cf. *(a)l* (sg) vs *i/j* (pl) in (12b)), whereas the third and sixth person of the verb are identical:

(12) a.   l'è   (da) tre ure   che ta   spete/so dre a spetat
           it=is (for) three hours that you= I.wait/have.been.waiting

     b.   a   j'è   (*da) tre ure che ta   spete
           a=it/they=is (for) three hours that you= I.have.been.waiting

Simple wh elements such as *quand(o)* 'when' and *quat* 'how long' always exhibit clitic inversion, even in the contexts in which inversion is normally banned, i.e. with past/future tenses, cf. (13c):

(13) a.   Quand é-l/*l è   che ta   egnet a troam?
           When is=it/it=is that you= come to meet=me

    b.   Quat   é-l/*a l è   che ta   lauret in svisera?
        How.long is=it/*a=it=is that you= work in Switzerland
    c.   Quat   ere-l/*a l era   che ta   lauraet in svisera?
        How.long was=it/*a=it=was that you= worked in Switzerland

Conversely, complex *wh* elements such as *quace agn* 'how many years' do not trigger inversion.

4.3 Other interrogatives

The remaining wh elements, e.g. *how, why,* etc. cannot occur in clefts. The following examples illustrate the contrast between CN and Italian:

(14) a. *Ndo   é-l   che l va?
      b.   Dov' è   che va?
         Where is=(it) that (he)=goes

(15) a. *Com' é-l   che al noda?
      b.   Com' è   che nuota?
         How is=(it) that he=swims

5. Summary and theoretical implications

The following table summarises the distribution of clefts in CN. Recall that, although clefts are permitted in declarative clauses, they do not yield a corrective interpretation and do not allow the focalisation of PPs.

(16) Focus:

Focus:	Declarative	wh Interrogative
DP	✓	✗
MP (temporal)	✓	✓
Others	✗	✗

Temporal clefts, by contrast, are always permitted, even in questions. It is worth noting that the behaviour of inversion sets apart temporal clefts from other types of clefts. The occurrence of inversion with any tense in combination with (simple) temporal wh elements points towards a separate analysis of temporal clefts. In a nutshell, we are going to argue that temporal clefts are in fact the only true clefts of CN, while other *prima facie* clefts are better analysed as concealed pseudo-clefts.

    Leaving temporal clefts aside, the data introduced so far challenge an analysis of clefts in terms of focus-movement. If clefting was a focalisation strategy, all kinds of XP would be expected to occur in clefts and no declarative/interrogative asymmetry should occur, *contra* evidence. To account for the observed restrictions, we

argue for an alternative analysis, according to which in some languages clefts and pseudo-clefts have the same structure (Paul 2001 a.o.; see also Percus 1997).

Let us assume that pseudo-clefts are equative copular constructions, i.e. a Small Clause (Heycock & Kroch 1999) headed by an equative head (=), which takes a headed Relative Clause as its complement.

(17) $[_{SC}$ DP = $[_{RC}$ DP …]]

The equative head establishes an identity relation between the subject of the SC and the head of the RC, e.g.

(18)   è $[_{SC}$ Luca = $[_{RC}$ quello che ho baciato ~~quello~~]]
     'Luca is the one that I kissed'

The equative relation holds if and only if the subject of the SC and the head of the RC have the same categorial features, thus ruling out cases like (19):

(19) $[_{SC}$ *PP = $[_{RC}$ DP …]]

This explains why PPs cannot occur in pseudo-clefts with a specificational reading (recall that, when locative PPs are allowed in a copular construction as in (20c), they can have only a locative interpretation, i.e. 'the place is *in* Milan', and not 'the Place *is* Milan', see fn 2):

(20)   a. *è   con Luca il   ragazzo con cui mi sposo
          it.ɪs with L.   the boy   that   I   marry
       b. *è   a matita   il   modo in cui   disegno
          it.ɪs with pencil the way   in which I draw
       c. *è   a Milano il   posto in   cui vado
          it.ɪs in Milan   the place where I   go

With this in mind, let us suppose that in CN clefts are concealed pseudo-clefts: they do not result from extraction from a (small) CP *à la* Belletti, see (1), but from an equative SC like (17). Under this analysis, we can easily account for the ungrammaticality of PP clefts, which in CN are as ungrammatical because of the constraint in (19) (the same holds for wh elements corresponding to PP arguments/adjuncts, cf. § 4.3).

By the same token, *prima facie* clefts such as (21) are therefore supposed to derive from a copular structure like (22), once the head of the RC (*chel* 'that') is deleted:

(21)   A l'è   ol Luca che l'ha mangiat la   turta
       a=it=is the L.   that he=ate   the cake

(22)   A l'è   [$_{SC}$ ol   Luca = [$_{RC}$ (chel) che l'ha mangiat la turta]]
       a=it=is   the L.   that   that ate the cake

To derive (21) from (22), one has to resort to a mechanism of deletion (not dissimilar from the one invoked for the analysis of relative clauses since Chomsky 1977), in which the head of the RC is deleted under identity with the subject of the SC (for an alternative machinery, see Percus 1997).

However, following this analysis, one wonders why deletion does not take place in interrogatives, where only pseudo-clefts are allowed (the relevant examples are repeated below for the sake of clarity).

(23)   a.   Chi é-l   *(chèl)   che l è   dre a durmì?   *who*-S
            who is=he   the.one that he=is sleeping
       b.   Chi é-l   *(chèl)   che ta   set dre ad ardà?   *who*-O
            who is=he   the.one that you= are watching
       c.   Cos é-l   *(chèl)   che ta   manget?   *what*-O
            what is=it   the.one that you= eat

We contend that deletion is blocked as the featural specifications of the equated XPs do not correspond: if the subject of the SC bears a [+ WH] (or Q) specification, we hypothesise that deletion cannot take place, thus giving rise to the observed asymmetry between declaratives, where pseudo-clefts are finally turned into *prima facie* clefts, and interrogatives, where the head of the RC cannot be omitted:

(24)   Chi é-l   [$_{SC}$ ~~chi~~ [$_{RC}$ *(chèl)   che l'è   dre a durmì]]?
       who is=he   the.ONE that he=is sleeping

The above solution might sound rather *ad hoc*. Nonetheless, it is worth noting that similar declarative/interrogative asymmetries are found in Italian as well and, to the best of our knowledge, have remained unnoticed so far. Let us start with the minimal pair in (25): the two sentences have the same meaning and the same information structure; in the former, the PP is extracted from the subordinate clause, whereas in the latter a DP element co-occurs along with an oblique wh element (*con cui* 'with whom').

(25)   a.   È   con Giorgio che voglio scappare
            it.IS with G.   that I want to escape
       b.   ˀÈ   Giorgio con cui   voglio scappare.
            it.IS G.   with whom I want to escape

Benincà (1978: fn. 2) points out that sentences like (25b) are slightly degraded and, instead of (25b), a pseudo-cleft such as (26) is normally preferred.

(26) È   Giorgio quello con cui   voglio scappare
     it.IS G.   the.one with whom I want to escape

Interestingly, the slight asymmetry noticed by Benincà becomes a full contrast once the sentences in (25) are turned into questions: the former remains grammatical, while the latter results in severe ungrammaticality.

(27) a. Con chi   è _ che vuoi   scappare?
         with whom is   that you.want to escape
     b. *Chi è _ con cui   vuoi   scappare?
         Who is   with whom you.want to escape

Again, (27b) is fine if we resort to a pseudo-cleft construction:

(28) Chi è _ quello   con cui   vuoi   scappare?
     Who is   the.ONE with whom you.want to escape

In the light of the analysis of CN, we argue that the sentences in (25) and (27) are not of the same kind: (25a) and (27a) are fully-fledged clefts, whereas (25b) and (27b) are concealed pseudo-clefts of the CN type. Underlyingly, they are copular constructions in which the head of the RC has been deleted:

(29) È [$_{SC}$ Giorgio = [$_{RC}$ ~~quello~~ con cui voglio scappare]]

Like in CN, the head of the RC cannot be deleted if the subject of the SC is interrogated:

(30) Chi è [$_{SC}$ _ = [$_{RC}$ *(quello) con cui voglio scappare]]

This is why only fully-fledged pseudo-clefts such as (25a) can occur in questions, while concealed pseudo-clefts such as (25b) cannot.

6.   Conclusions

This contribution has focused on the syntax of clefts in the Lombard dialect of Comun Nuovo (Bergamo). We have shown that in this dialect clefts are highly constrained and, in the contexts where clefts and pseudo-clefts alternate, the distinction between the two is often blurred.

In particular, clefts are ungrammatical in questions, save for temporal clefts which, however, exhibit a puzzling behaviour with respect to inversion (inversion normally occurs with certain types of wh elements and in sentences with a deictic present tense).

Leaving temporal clefts aside, we contend that CN clefts are better analysed as concealed pseudo-clefts, an analysis put forth for other non-European languages (see Paul 2001 and references therein). We do not claim that the usual raising analysis of clefts must be always replaced by a pseudo-cleft analysis, but we point out that the divide between clefts and pseudo-clefts is far less straightforward than usually thought and that the boundary between the two analyses is ultimately an empirical matter.

Acknowledgments

References

Belletti, Adriana. 2008. The CP of clefts. *Rivista di Grammatica Generativa* 33: 191–204

Benincà, Paola. 1978. Sono tre ore che ti aspetto. *Rivista di Grammatica Generativa* 3(2): 321–345

Benincà, Paola. 1983. Il clitico a nel dialetto padovano. In *Scritti Linguistici in Onore di Giovan Battista Pellegrini*, Paola Benincà et al. (eds), 25–32. Pisa: Pacini. Reprinted in Beninca 1994, *La variazione sintattica*. Bologna: il Mulino.

Chomsky, Noam. 1977. *On wh-movement*. In Formal syntax, P. Culicover, T. Wasow, and A. Akmajian (eds), 71–132. New York: Academic Press.

Collins, Peter C. 1991. *Cleft and Pseudo-cleft Constructions in English*. London: Routledge.

Donzelli, Giulia. 2017. *Standard and special questions in Comunuovese*. Talk given at *Romance Interrogatives. Workshop for doctoral students on question forms in Romance Language*, Universität Konstanz, 18–19 May.

Heycock, Caroline & Kroch, Anthony. 1999. Pseudocleft connectedness. Implications for the LF interface level. *Linguistic Inquiry* 30: 365–397.

Manzini, M. Rita. 2014. Grammatical categories: Strong and weak pronouns in Romance. *Lingua* 150: 171–201.

Manzini, M. Rita & Leonardo M. Savoia 2011. Wh-in situ & wh-doubling in Northern Italian varieties: Against remnant movement. *Linguistic Analysis* 37: 79–113.

Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111: 707–727

Percus, Orin. 1997. Prying open the cleft. In *Proceedings of NELS* 27, Kiyomi Kusumoto.(ed.), 337–351. Amherst MA: GLSA.

Parry, Mair. 1997. Variazione sintattica nelle strutture interrogative piemontesi. In *Strutture interrogative dell'Italia settentrionale*, Paola Benincà & Cecilia Poletto (eds). *Quaderni di Lavoro ASIS* 1: 91–104.

Poletto, Cecilia. 2000. *The Higher Functional Field.Evidence from Northern Italian Dialects*. OUP: Oxford.

Poletto, Cecilia & Vanelli, Laura. 1997. Gli introduttori delle frasi interrogative nei dialetti italiani settentrionali. *Quaderni di Lavoro ASIS* 1: 105–117.

Poletto, Cecilia & Pollock, Jean-Yves. 2006. Another look at wh-questions in Romance. The case of mendrisiotto and its consequences for the analysis of French wh-in situ and embedded interrogative. In *Romance Languages and Linguistic Theory* [Current Issues in Linguistic Theory 303], Danièle Torck & W. Leo Wetzels (eds), 199–258. Amsterdam: John Benjamins.

Salvi, Giampaolo. 1991. Le frasi copulative. In *Grande grammatica italiana di consultazione*, Vol. II, Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds), 163–189. Bologna: il Mulino.

Sornicola, Rosanna. 1988. It-clefts and wh-clefts: Two awkward sentence types. *Linguistics* 24: 343–379.

Valentini, Ada. 2012. Per una tipologia della struttura informativa: Il caso delle frasi scisse in un dialetto italoromanzo. *Linguistica e Filologia* 32: 75–117.

CHAPTER 9

Negation patterns across dialects

Cecilia Poletto and Michèle Oliviéri

Goethe University Frankfurt / University of Padua / University of Côte d'Azur, CNRS, BCL

In this work we consider two opposite sorts of Romance varieties with respect to the negative marker, i.e. an Occitan variety where the preverbal negative marker alone is so much reduced that it can nowadays only occur in some specific syntactic contexts and two Venetan varieties where the preverbal negative marker is so widespread that it even displays strict negative concord. We will show that despite being so different, both types of dialects are sensitive to the modal environment and that the presence of the preverbal negative marker is tied to a subset of non-veridical contexts, i.e. those that contain a [−realis] verbal form.

Keywords: negation, Occitan dialects, Northern Italian dialects, modality

1. Introduction

The Jespersen cycle and the way negative markers are renewed is an ever green topic of French syntax and from the seminal work by Zanuttini (1997) also of Northern Italian dialects. Lot of work has concentrated on the way the new negative marker establishes itself, in which contexts it occurs first and what properties it displays in relation to other negative elements in the clause. In this work we rather concentrate on the old negative marker, which in Romance is generally represented by the alveolar nasal plus a vowel n + V+(n). The research question we intend to answer is the following one: can we identify common factors that enhance the presence of the preverbal negative marker across Romance? The hypothesis we will put forth is that modality is at least one of the factors that favor the presence of the preverbal negative marker in varieties that are losing the preverbal negative marker but also in those which use it more than standard Italian or Spanish. This means that among the factors influencing the Jespersen cycle (JC), we have to include modality, a proposal that to our knowledge has never been made so far. In order to do so, we will compare two dialects that are at the opposite poles of the JC spectrum, namely the Eastern Occitan variety of L'Escarène in the Nice hinterland

(France) and two North Eastern Italian varieties, the one spoken in Zemignana, at the borders of the provinces of Padua and Venice and Venetian.

While in L'Escarène the preverbal negative marker is progressively disappearing and is maintained in a limited number of contexts, in Zemignana and Venice it is very well preserved. They even extend the usage of the preverbal negative marker and give origin to the so-called phenomenon of strict negative concord (we use here the definition by Zeijlstra 2004). In other words, the negative marker occurs both with pre- and postverbal n-words, differently from standard Italian, which is a non-strict negative concord language.

Despite being at the two extremes of the spectrum of the life span of the preverbal negative marker, the Occitan and Venetan varieties considered display some surprising similarities, which can lead us towards interesting observations on the syntactic interaction between negation and modality. As we will see, in both dialects the preverbal negative marker appears more frequently in modal contexts. We will show that the abstract common feature that has an impact on the distribution of the negative marker is the same in both cases and has to do with the [−realis] status of the construction/context.

2. The data

2.1 L'Escarène

The first interesting observation that comes from the data gathered on Escareasc is that this dialect displays at the same time all the three stages of the Jespersen's cycle.[1] The first one is the occurrence of a preverbal negative marker alone (*noun*), the second one shows a combination of the preverbal marker with a postverbal element (typically *pa*) and at the third one, only the postverbal marker emerges. Although there is a part of variation, these three types of negation correspond to different structures.

2.1.1   *Single postverbal marker*

Negation is more frequently expressed with a single postverbal element, the unmarked *pa*, or another n-word such as *plu, rèn, jamai, degun*, and this structure emerges mainly in independent or main clauses, so with the indicative (1).

(1)   a.   Sabi   pa.
            I.know not
            ´I do not know´.

       b.   Parti   pa souvèn.
            I.leave not often
            ´I do not go away very often.´

       c.   Sènte   bouòn, troves   pa?
            It.smells good,   you.think not
            ´It smells good, don't you think?´.

       d.   I fèn   pu   atencioun.
            it we.do no.more attention
            ´We no longer pay attention to it.´

       e.   Fèii rèn   couma lis autres.
            I.did nothing like   the others
            ´I did not do anything like other people.´

       f.   Si   sau   jamai.
            REFL know never
            ´You never know.´

       g.   Veirès   pu   degun.
            YOU.see.FUT no.more nobody
            ´You will not see anybody anymore.´

       h.   Si   pouor pa faire beure u ase   que a   pa set.
            REFL can   not make drink a donkey that has not thirst
            ´You can lead a horse to water but you can't make it drink.´

       i.   Voudrìi   pa que m'=en=vourguesses.
            I.want.COND not that to.ME=of.IT=you.want.SBJV.IPFV
            ´I would not want you to hold it against me.´

       j.   Cresi pa que vouoron i   anar.
            I.think not that they.want there go
            ´I do not think that they want to go there.´

It is also found in embedded indicative or infinitive clauses (2), but less frequently with other N-words than *pa*.

(2)   a.   Siou sugur que ti   sies   pas assè   pausà.
            I.am sure   that you.REFL you.are not enough rested
            ´I am sure that you did not rest enough.´

    b.   Pensi   que   as   pa rasoun.
         I.think that you.have not right
         ´I think that you are not right.´

    c.   M=al   dich que pouia   pa lou faire perqué arribava   pa a
         to.ME=he.HAS said that he.could not it   do   because he.can.PST not to
         s'=endurmir   bouòn' oura.
         REFL=fall.asleep good   hour
         ´He told me that he could not do it because he was not able to fall asleep
         early.´

    d.   Ai   paur de pa saupre mi desbroulhar.
         I.have fear of not know me manage
         ´I am afraid that I will not know how to manage.´

    e.   Es l'istòria d'un ase   qu'a   jamai quità lou siou païs.
         It.is the story of a donkey that he.has never left   the his village
         ´It is the story of a donkey who never left his village.´

    f.   Dèia   sèmpre que pouria   faire rèn.
         he.said always that he.could do   nothing
         ´He would always say that he could not do anything.´

    g.   Fa   quinze ans qu' èri plus   vengù.
         it.makes fifteen years that I.was no.more come
         ´I have not come for fifteen years.´

    h.   Pènsi que troveras   degun per t'=ajuar
         I.think that you.find.FUT nobody for you=help
         ´I think that you will not find anyone to help you.´

As we will see the contexts in which the preverbal negative marker can occur alone
are quite different.

2.1.2   *Single preverbal noun*

– *imperative clauses.*   The preverbal negative marker *noun* obligatorily occurs
alone with the imperative, in the singular (expressed by the infinitive) and in the
plural (expressed by the subjunctive).

  (3) a.   Parle=mi!
         talk.IMP.SG=to.me
         ´Talk to me!´ (singular)

     a′. Parlà=mi!
         talk.IMP.PL=to.me
         ´Talk to me!´ (plural).

     b.   Noun mi=parlar!
         NEG   to.me=talk.INF
         ´Don't talk to me!´ (singular)

b′.   Noun mi=parlès!
  NEG   to.me=talk.SBJV
  ´Don't talk to me!´ (plural).

As the examples in (3) show, the form of the negative imperative (3b), (3b′) is not the same found in positive imperatives (3a), (3a′), and the singular corresponds to an infinitival form. This type of phenomenon has been reported and analyzed by Zanuttini (1997) for Italian (see below). In the plural, the verbal form is the subjunctive one and also displays the preverbal marker *noun*.

2.1.2.2   *embedded clauses.*   The other typical context in which the preverbal negative marker occurs, once again without any postverbal negative marker, are subjunctive and conditional embedded clauses, whether it is a completive (4) or a circumstantial clause (5).[2]

(4)   a.   Voudrìi   que noun plòuguesse.
         I.would.like that NEG   rain.SBJV.IPFV
         ´I would like it not to rain.´

      b.   Voudrìi   que noun partesses.
         I.would.like that NEG   you.leave.SBJV.IPFV
         ´I would like you not to go.´

      c.   Pensi   que noun siguès   embilaia.
         I.think that NEG   you.are.SBJV.PRS angry.F
         ´I think that you are not angry.´

      d.   Carìa   que noun lou trouvès.
         it.must.be.COND that NEG   it   he.finds.SBJV.IPFV
         ´It would be necessary for him not to find it.´

      e.   Titoun es d'avis   que noun si=pourria   faire   pu   ben.
         Titoun is of.opinion that NEG   REFL=can.COND do.INF more well
         ´Titoun thinks that that it could not be done better.´

(5)   a.   Scounde acò, que noun lou=trove.
         hide   this, that NEG   it=find.SBJV.PRS
         ´Hide that, so that he cannot find it.´

      b.   Era   stacà   en un piqué per que noun pousquesse   courre en
         It.was hitched in a   post for that NEG it.can.SBJV.IPFV run   in
         lou prà.
         the meadow
         ´It was hitched to a post so that it could not run in the meadow.´

---

2. The corpus does not allow to test the relative clauses, more investigation is still needed, but we expect the same behaviour in relative clauses.

    c.   Se noun m'=eloigni,   pilhi pa un dangier.
          If NEG me=i.walk.away.IND, i.take not a risk
          ´If I do not walk away, I do not take a risk.´

These observations strongly suggest that there is a correlation between the presence of the negative marker *noun* and the mood of the verb. Considering that imperative and subjunctive share an aspectual [−realis] meaning/feature, this feature could trigger the occurrence of the preverbal *noun*, here without any postverbal negative marker. Another piece of evidence in favor of this analysis is the presence of the preverbal *noun* in if-clauses (5c), by definition [−realis], despite the absence of the subjunctive mood.

    At this point, a kind of complementary distribution appears between these two negative markers, the postverbal marker emerging in a [+realis] clause, while the preverbal marker is devoted to [−realis] interpretations. This corresponds to the situation of the dialect spoken in Nice (Nissart), which represents a later stage of the diachronic evolution of negation.

2.1.3 *Preverbal noun and postverbal marker*

In Escareasc, which is more archaic than Nissart, a third type of negation combines this preverbal element *noun* and a postverbal marker. It occurs sometimes in independent/main clauses (6), and more frequently in embedded clauses whatever the mood, subjunctive (7) or indicative/Infinitive (8):

(6)   a.   Lis enfans,   noun lis   ai   pa vist.
          the children, NEG   them i.have not seen
          ´I have not seen the children.´

        b.   N'=aimes   pa quarqu'un?
          NEG=you.like not someone
          ´Do you not like someone?´

        c.   N'=avìa   rèn   audì ni degun.
          NEG=he.HAS nothing heard nor nobody
          ´He had not heard anything or anyone.´

        d.   Noun sabi   plus!
          NEG   I.know anymore
          ´I do not know anymore.´

        e.   Li cuòl(s) de Nissa e   de Castilhoun noun soun pa tròu aut(s).
          the passes of Nice and of Castillon   NEG are   not too high
          ´The Nice and Castillon passes are not too high.´

(7)   a.   Carìa   que noun plòuguesse   pu.
          it.must.be.COND that NEG   it.rains.SBJV.IPFV anymore
          ´It would be necessary for it not to rain anymore.´

b. Mi=farìa tout plen plesir que noun
to.ME=it.makes.COND all full pleasure that NEG
lou=desias pa.
it=you.say.SBJV.PRS not
´I would be glad if you did not say it.´

c. A paour que noun sabés rèn.
she.has fear that NEG you.know.SBJV.PRS nothing
´She is afraid that you do not know anything.´

d. Esperèn que noun partés pa encara.
we.hope that NEG you.leave.SBJV.PRS not yet
´We hope that you are not leaving yet.´

e. Vouor que noun digés rèn.
he.wants that NEG you.say.SBJV.PRS nothing
´He wants you to say nothing.´

f. Quèn que noun regardés degun.
it.must.be that NEG you.look.SBJV.PRS nobody
´You must look at no one.´

(8) a. Si=crese que noun sabi rèn.
REFL=she.believes that NEG I.know nothing
´She believes that I do not know anything.´

b. Cresi pa que noun parles pa.
I.believe not that NEG you.talk not
´I do not believe that you do not talk.´

c. A paour que noun parles pa.
she.has fear that NEG you.talk not
´She is afraid that you will not talk.´

However, except in subjunctive embedded clauses (obligatory displaying *noun*
alone), this structure seems to alternate freely with the former one (9).

(9) a. Noun sabi pa cenque fon en la journaia. Sabi pa cenque
NEG I.know not what they.do in the day I.know not what
fon en la journaia.
they.DO in the day
´I do not know what they do during the day.´

b. Ma noun vourìi pa lou=laissar enfermà touta la semaa. Ma
but NEG I.want.PST not it=leave.INF locked all the week but
aurii pa vourgù lou=laissar serrar touta la semaa.
I.have.PST not wanted it=leave.INF locked all the week
´But I did not want to leave him shut up inside all week long.´

  c. Noun deu pa èstre tardi. Deu pa èstre ton tardi.
    NEG it.must not be.INF late it.must not be.INF very late
    ´It must not be (very) late.´

  d. Noun cresi pa que vouòrgon   veïr.
    NEG I.think not that they.want.SJBV.PRS come.INF
    Cresi pa que vouòrgoun   veïr.
    I.think not that they.want.SJBV.PRS come.INF
    ´I do not think that they want to come.´

Although the informants claim that there is no semantic-pragmatic difference in these pairs, one can wonder if it is not a remnant of an older system where different aspects were distinguished, as it has been reported by Camproux (1958) or Lafont (1967).[3] Even if the speakers are not aware of this fact, the presence of *noun* here could actually manifest an aspectual value.[4]

  We conclude our brief tour of negation by pointing out that the preverbal negative marker can only occur alone (a) with negative imperatives, (b) with [−realis] forms (present and past subjunctive, future indicative, conditional). We will now see how these two contexts are also special in Venetan varieties.

2.2 Venice

In Venetian, and more generally in Venetan dialects, negative imperatives are either marked through a different type of negative marker, which is not the head / no/ with a closed vowel, but /nɔ/ with an open vowel, which is also the form of pro-sentence negation:

 (10) Ti vien?
    You come?
    'Are you coming?
    Nɔ.
    no
    ´No, I am not.´

---

3. For a detailed presentation, see Oliviéri et Sauzet (2016: 348–349).

4. We also have to notice that our corpus contains some exceptions, although very few, such as the following ones:
 Parles pa. / Noun parles. ´You do not talk´
 Noun plouravi. / Plouravi pa. ´I was not crying.´
 A par acò, noun vouòri de bèn en degun ´Otherwise, I do not like anybody.´
 Perqué noun travalhes? ´Why don't you work?´
 Se counouisses degun, ti farai rescountrar de gèns d'aicì. ´If you do not know anybody, I will introduce you to people from here.´
 Noun mi plas que degun noun m'invite jamai. ´I don't like it that nobody ever invites me.´

(11) nɔ   tocar, sa!
      NEG touch, PARTICLE
      ´Do not touch!´

(12) nɔ   dirghe cossí.
      NEG tell.him so
      ´Do not tell him so´.

Interestingly, the "normal" preverbal negative marker with a closed vowel is only possible if the auxiliary *sta*, a reduced form of the verb ´stay´, ´remain´ is added, as noted by Kayne (1992), who analyzes this as a modal auxiliary:

(13) No sta   tocar, sa!
      NEG MOD touch, PARTICLE
      ´Do not touch!`
      No sta   dirghe cussí.
      NEG MOD. tell.HIM SO
      ´Do not tell him so´.

One might wonder whether the pro-sentential form of the negative marker is actually a substitute for both negation and the modal auxiliary. If we compare Venetian and Escareasc, we immediately notice that in both languages imperatives require a special type of negative marker (or in Venetan an additional modal auxiliary). Zanuttini (1997) puts forth the generalization that those Romance languages that have a preverbal negative marker must substitute true imperative forms with suppletive forms.[5] Zanuttini's idea is that imperative forms lack the Tense projection, which is required by the preverbal negative marker, and therefore must be substituted through a form which projects it, otherwise the NegP projection cannot select TP as its complement. This analysis readily explains why the verbal form must change when a preverbal negative marker is present, but does not account for cases in which it is the negative marker that changes and not only the verbal form. Venetian and Occitan are parallel in the sense that imperatives require a special type of negative marker (or alternatively Venetian also has a different repair strategy as the modal auxiliary, and in this case it can use the standard negative marker, as noted above).

The second special case found in L'Escarène has to do with [−realis] verbal forms, and there does not seem to be any parallel between Venetian and Escareasc in this case, since Venetian always has the preverbal negative marker with all types of verbal inflection. However, there are some hints that something like modality is

---

5. She defines a true imperative as a form which is unambiguously only imperative and not homophonous with any other verbal form.

relevant also for negative concord in Venetan as well. To show this, we will present data from Zemignanese, another Venetan dialect spoken in the mainland about 30 kilometers from Venice.

2.3   Zemignana

As mentioned above, Zemignanese, on a par with other Venetan dialects of the Venice province, allows for strict negative concord. In Zemignanese negative concord is less frequent than in the Venetian variety we have examined above, and in particular strict negative concord seems to be limited in embedded clauses by the type of main verb. Here we report an investigation made in Solivo (2017) on 30 native speakers through a judgement test: subjects were required to place the same sentence with and without negative concord on a scale from 1 to 5.[6]

(14)   Credea   che nianca   (no)l vegnese casa   magnare.
       Thought that not.even not he came   home eat.INF
       ´I thought that he would not even come home to dinner´.

(15)   Me incorzo sempre tardi che nianca   (no)l ghe ze   el Sabo
       Me notice   always late   that not.even (not) he   there is the Saturday
       de sera.
       of night
       ´I always notice too late that he does not even come home on Saturday night´.

In general, both variants, with and without strict negative concord are possible, although differences emerge in the judgements. The results are provided in the following tables:

The first table is referred to an embedded clause under the verb *credere* ´believe´, which requires the subjunctive form in both positive and negative clauses. The second table referred to an embedded clause under the verb *incorzarse* ´to notice´, which requires an indicative form on the embedded verb. We can conclude that one of the factors favoring the occurrence of the additional negative marker in strict negative concord contexts is modality.

---

6. The judgement task included 50% of fillers and two conditions: mood selection on the embedded verb and main verbs that imply the truth of their complement. These two conditions were tested for the following n-words *gnanca* 'not even' *neancora* 'not yet' *nisuni* 'nobody' and *ninte* 'nothing'. In total there were 48 task sentences administered to the subjects in one session.
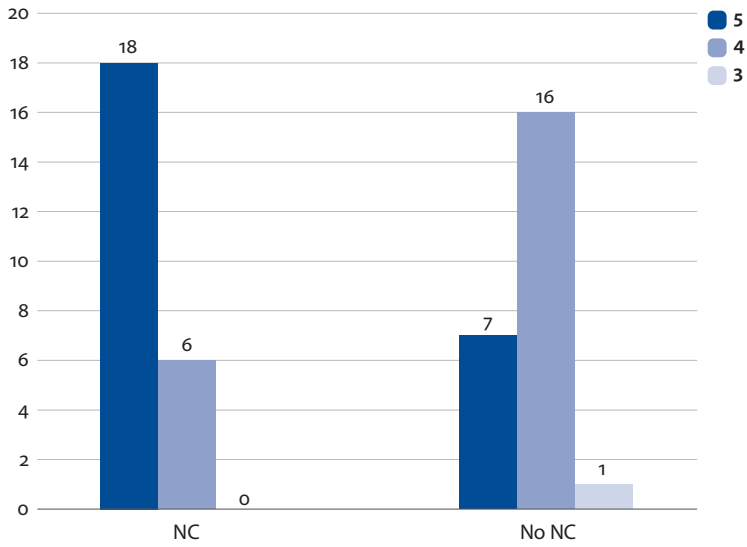
Figure 1. Negative concord with the main verb *credare* 'believe' and the adverb *nianca* 'not even'
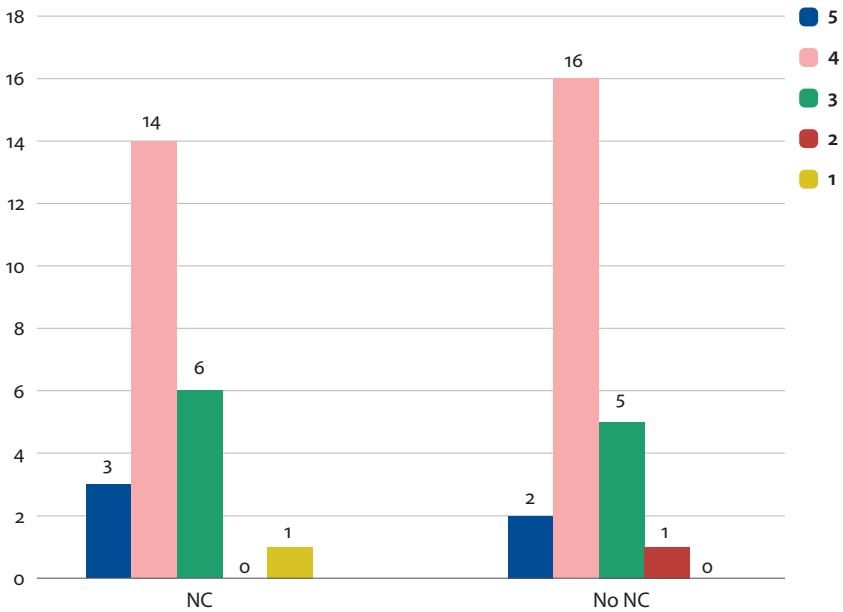

Figure 2. Strict NC with the verb *incorsarse* 'to notice' and adverb *nianca* 'not even'

3.   Putting the pieces together: Negation and modality

Summing up what we have seen so far, we see that in the Occitan dialect of L'Escarène there are two contexts in which the preverbal negative marker is still preserved and can appear alone: (a) the first is the case in which the inflected verb corresponds to a subjunctive or conditional form, i.e. a [−realis] for, (b) the second is the negative imperative context, which is also [−realis].

In Venetian we also find something similar to (b), i.e. as noticed by Zanuttini (1997) real imperatives and the standard preverbal negative marker are incompatible. In Venetan, negative imperatives use a different type of negative marker (in this case the pro-sentence negation) or another repair strategy with a modal auxiliary. Also Venetan dialects display sensitivity to other [−realis] verbal forms, as Zemignanese clearly shows for the extension of strict negative concord to modal environments. At this point, we wonder if there is a way of unifying these contexts in both languages. If we consider subjunctive and imperatives as two different forms of modality which both express a sort of [−realis] macrofeature, then we have the parallel we were looking for. This has the advantage that we can have a unified analysis for all contexts in both L'Escarène and Venetan: the preverbal negative marker tends to occur in modal [−realis] contexts.

This in turn means that in order to tie preverbal negation to modality in negative imperatives we need to reformulate Zanuttini's proposal in the following way: it is not the lack of Tense that blocks the co-occurrence between the preverbal negative marker with true imperative forms. It is rather the necessity to have a modal agreement between [−realis] modality and negation and this agreement can either be morphologically expressed by changing the verb into a different form or by changing the negative marker into a different form or both. More specifically, we would like to adopt Kayne's (1992) idea that the preverbal negative marker licenses a null modal element and this is the reason why in negative imperatives clitics can occur in proclisis, while in positive imperatives they only occur in enclisis even in languages like French, which do not have any other context where object clitics are enclitics. Kayne provides several arguments in favor of the presence of a null modal in negative imperatives, one of which we have already considered, namely the presence of a lexical modal auxiliary in Venetian (see above (13)). Since in Venetian the modal auxiliary is visible, it is difficult to deny its existence. The basic idea we put forth is thus that the different form the negative marker takes in negative imperatives is an agreeing form which is able to license the null modal auxiliary. There are several languages that display a richer morphology on the element licensing a null category, like for instance Celtic languages, where null subjects are related to a stronger form of agreement on the verb. The idea to apply Kayne's hypothesis concerning the presence of a modal in

negative imperatives thus explains not only (a) the presence of a lexically realized modal form in Venetian (b) the infinitival form of the lexical verb (c) the possibility of proclisis, but also (d) the different morphology of the negative marker, which is different precisely because it has to license the null modal. The reason why negation can license a modal can be seen as an effect of a sharing of the [−realis] features, or in semantic terms by the fact that both negation and modals are non-veridical, as Giannakidou (1998) pointed out.[7] The licensing of null elements through a procedure of agreement is actually rather standard (pro drop is but an example of this) so that the fact that negation can license a null modal while other elements do not is not surprising.

This means that languages like Greek, which have special forms for the negative marker in modal contexts are similar to those Romance languages where we see that imperative forms change into something else when the negative marker is present. There are other Romance varieties where we see the same effect, i.e. negative imperatives do not change the form of the verb but rather the form of the negative marker. One such case is the Rhaetoromance dialect of San Leonardo examined in Poletto and Zanuttini (2004), where the usual discontinuous negative marker *ne…nia* is substituted by the negative marker corresponding to pro-sentence negation (like in Venetian) *no*, which can occur pre- or postverbally:

(16) Maria ne vegn nia/mine a ciasa.
      'Maria isn't coming home.'
      *Ne le fa nia/mine!
       Not it do not
      'Do not do it!'
      *Nia/mine le fa!

(17) Ne le fa no!
      'Do not do it'.
      No le fa!

Notice that here it is not the verbal form which changes into an infinitive, subjunctive or gerund, as it is the case in other Romance varieties, it is the negative marker that changes, so that the postverbal part of the discontinuous negation is substituted by the pro-sentence negative marker, which can also occur preverbally cancelling the preverbal morpheme. Furthermore, as expected under Kayne's analysis, the clitic can only be proclitic in these cases.

The idea that the change in the verbal form in negative imperatives is the counterpart of the change in the form of the negative marker and that both do not

---

7. This predicts that other non-veridical contexts like wh-interrogatives should also display the same pattern. At present, we are testing this for Venetan.

have to do with the lack of a Tense projection, but are rather the result of a null modal being licensed has the advantage to cover both phenomena under the same explanation. If we assume that the [−realis] feature is actually what Giannakidou refers to as non-veridical contexts, we actually expect them to have common properties, and that in some languages this is expressed through morphology. If we are correct in our view, the presence of the preverbal negative marker is sensitive to non-veridical contexts both in Romance languages that are losing it and in varieties that have extended it to strict negative concord, since this derives from a general Agree procedure. At this point one might wonder why postverbal negation does not agree (and license) the null modal, but actually there are sporadic cases in which the postverbal negative marker licenses the null modal, so that we see an infinitival form of the lexical verb depending on the modal. This has already been noticed in Benincà and Poletto (2004) for Emilian dialects where the lexical verb has an infinitival form, but the negative marker is the postverbal *mia*[8]:

(18) Movrat   mia.
     move.INF.yourself not
     ´Do not move`.

One might wonder why it is the case that postverbal negation licenses the null modal only in few dialects, while the preverbal negative marker invariably does so. If we make a parallel to the most well-known case of agreement, namely subject-verb agreement, we notice that V to T is in general a pre-requisite for the licensing of null subjects in SpecT. This means that since the preverbal negative marker, which is structurally close to the null modal can license it, while postverbal negation, which is not structurally close to the modal does so only in a far more restricted set of languages.

Summing up: assuming that the preverbal negative marker and the modal auxiliary agree for the [−realis] feature explains the following facts (a) why we find alternation between lexical and null modals (exactly as we find alternation between pro and lexical subjects) (b) why sometimes the negative marker changes its form, since the different morphology is the effect of the agreement procedure (c) why the infinitival form, i.e. the null modal selecting it is found mainly with the preverbal negative marker and only rarely with the postverbal one, since Agree plus movement licenses null elements more frequently than Agree by itself (see Guasti & Rizzi 2002 for a discussion on this).

---

8. Notice that here the clitic is in enclisis, but this is actually expected in the NIDs, since they generally have no clitic climbing even with lexical modals. Notice furthermore that these dialects have no preverbal negative marker.

Furthermore, it also explains why we find the preverbal negative marker with imperatives and [−realis] verbal forms in both varieties that are losing the preverbal negative marker and varieties in which the preverbal negative marker is stable.

4. Conclusion

In this work, we have examined two opposed sets of dialects: on the one hand an Occitan variety, which is losing the preverbal negative marker and on the other hand two Venetan dialects, where the preverbal negative marker occurs in contexts that are generally banned from standard Romance languages like Italian and Spanish. We have noticed that there are some surprising symmetries in the occurrence of the preverbal negative marker, which allow us to establish a natural class of constructions, namely [−realis] verbal forms including subjunctive, conditional, future indicative and negative imperatives. We have put forth an explanation in terms of agreement between non-veridical operators which share a [−realis] feature through a standard Agree operation. Since negative concord is seen as an agreement procedure, it is also plausible to think that other types of non-veridical operators like modals are subject to a similar procedure of Agree. This opens up the possibility to capture two sets of phenomena: (a) the change in the verbal form in negative imperatives and (b) the change in the form of the preverbal negative marker found in some dialects. Furthermore, it also explains why cases of null modals are also sporadically found with postverbal negative markers, though this is not the standard case as it is with the preverbal negative marker: this asymmetry is the same found with Agree+copy (i.e. movement) with respect to Agree without any further operation. If this idea proves to be correct for other languages, it might lead to consider modality as one of the factors involved in the JC, which is either generally considered of phonological origin, as Jespersen himself first hypothesized, or can depend on some pragmatic values (like focusing negation, as put forth by van der Auwera in various recent papers, see van der Auwera 2010; van Alsenoy & van der Auwera 2014). For future research, we leave the problem of the position of the preverbal negative marker, which might be related to features located in the left periphery of the clause, either through verb movement (as in the case of negative imperatives) or by selection (as in if-clauses, or embedded subjunctives).

References

Benincà, Paola & Poletto, Cecilia. 2004. Topic, Focus and V2: defining the CP sublayers. In *The structure of CP and IP*, Luigi Rizzi (ed.) 52–75. Oxford: OUP.

Camproux, Charles. 1958. *Etude syntaxique des parlers gévaudanais*. Paris: Presses Universitaires de France.

Dalbera, Jean-Philippe et al. 1992. *Thesaurus Occitan (Thesoc)*. Université Côte d'Azur, CNRS, BCL, France. <http://thesaurus.unice.fr.> (22 November 2017).

Giannakidou, Anastasia 1998, *Polarity Sensitivity as (Non)Veridical Dependency*, John Benjamins, Amsterdam.

Giannakidou, Anastasia. 1999. *Polarity Sensitivity as (Non)Veridical Dependency*.[Linguistik Aktuell/Linguistics Today 23] Amsterdam: John Benjamins.

Guasti, Maria Teresa & Rizzi, Luigi. 2002. Agreement and tense as distinct syntactic Positions: Evidence from acquisition. In *Functional Structure in DP and IP*, Guglielmo Cinque (ed.) 167–194. Oxford: OUP.

Kayne, Richard S. 1992. Italian negative infinitival imperatives and clitic climbing. In *Hommages à Nicolas Ruwet*, Liliane Tasmowski & Anne Zribi-Hertz (eds) 300–312. Ghent: Communication & Cognition.

Kayne, Richard S. 2000. *Parameters and Universals*. Oxford: OUP.

Lafont, Robert. 1967. *La phrase occitane*. Paris: Presses Universitaires de France.

Oliviéri, Michèle & Sauzet, Patrick. 2016. Southern Gallo-Romance (Occitan). In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 319–349. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677108.003.0019

Poletto, Cecilia & Zanuttini, Raffaella. 2004. Making Imperatives: Evidence from Central Rhaetoromance. In *The Syntax of Italian Dialects*, Christina Tortora (ed.) 175–206. Oxford: OUP.

Solivo, Giulia. 2017. *The Phenomenon of Strict Negative Concord in Zemignaese*. MA thesis, University of Padua.

Van Alsenoy, Lauren & van der Auwera, Johan. 2014. On the relation between double clausal negation and negative concord. In *The Diachrony of Negation*, [Studies in Language Companion Series 160] Maj-Britt Mosegaard Hansen & Jacqueline Visconti (eds) 13–46. Amsterdam: John Benjamins.

van der Auwera, Johan. 2010. On the diachrony of negation. In *The Expression of Negation*, Laurence R. Horn (eds) 73–109. Berlin: Mouton De Gruyter.

Zanuttini, Raffaella. 1997. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford: OUP.

Zeijlstra, Hedde. 2004. *Sentential Negation and Negative Concord*. PhD dissertation, University of Amsterdam. Utrecht: LOT Publications.

A note on left-peripheral maps and interface properties

Luigi Rizzi

University of Genève & University of Siena

The cartographic analysis of the left periphery leads to the identification of invariant and variable properties in the syntactic expression of scope-discourse configurations, such as topic-comment and focus-presupposition. One notable property is that languages typically permit a unique focus in the left periphery of a clause, whereas left-peripheral topics may proliferate in many languages. A comparative analysis of Italian and Gungbe reveals that Italian disallows distinct LP foci also in distinct clauses of complex sentences, an option which is permitted in Gungbe. The proposal developed in this paper capitalizes on computational mechanisms applying at the interfaces with sound and meaning to capture the invariant core and the variability in these left-peripheral properties across languages.

Keywords: cartography, left periphery, focus, interfaces

1. Introduction

The cartographic enterprise involves a large descriptive and comparative dimension: in the first place, we want to know what the right structural maps are for the different zones of the syntactic tree, and what kinds of invariant and variable properties of the maps we can find across languages. This dimension has shown a strong heuristic capacity, leading to much work on numerous languages, and to the identification of cross-linguistic generalizations. Nevertheless, drawing the empirically correct maps is only the first step of the research. Questions of explanation immediately arise: why do we find the invariant properties that are observed across maps? And what is the appropriate parametrization to capture the cross-linguistic variation?

Addressing these issues immediately implicates a theoretical dimension: as soon as we ask questions on the deeper reasons of the arrangements we observe,

cartography can become a powerful generator of empirical issues for syntactic theory, which can enlarge the empirical basis of theoretical studies. See Cinque & Rizzi (2010), Rizzi & Cinque (2016), Rizzi (2017) for discussion.

After a quick presentation of the criterial approach to the syntactic expression of scope-discourse properties, I will discuss an empirical generalization involving topic and focus structures: while topics may proliferate, at most one left peripheral-focus seems to be allowed per clause. This constraint seems to hold in general, but languages clearly differ in the possibility of allowing distinct left-peripheral foci in distinct clauses of the same complex sentences: Italian disallows such a long-distance co-occurrence, while Gungbe permits it. In this note, I will try to capitalize on the computational mechanisms applying at the interface of syntactic representations with systems of sound and meaning to capture the observed invariant and variable properties in this important area of A-bar syntax.

2.   Syntactic and interface properties of criterial heads

In the cartographic study of the left periphery of the clause, the criterial approach to constructions expressing scope-discourse properties has played a crucial role (Rizzi 1997 and much subsequent work). According to this approach, the complementizer system is in fact a complex structural zone populated by a sequence of functional heads (Top, Foc, Q, Rel, Excl, …) which have a dual function:

–   In syntax, they trigger movement, attracting to their specifier a phrase endowed with matching features. So, a Q head attracts a phrase endowed with a + Q feature, a Top head attracts a phrase endowed with a + Top feature, etc.
–   At the interfaces with sound and meaning, criterial heads trigger interpretive procedures for the proper assignment of scope-discourse properties at LF, and the appropriate intonational contour at PF.

So, A'-constructions expressing scope discourse properties like topicalization or focalization involve representations such as the following, with criterial heads in bold:

(1)   a.   This book   **Top**   you should read _ tomorrow
       b.   THIS BOOK   **Foc**   you should read _, not Bill's book

Top and Foc are not pronounced in English, but this structure-based approach is made plausible by the existence of many languages in which criterial heads are overtly realized, e.g. the languages with overt topic and focus markers such as the African language Gungbe (cf. Aboh 2004):

(2) a.   Un sè   [do [dan lo **yà** [Kofi hu   ì]]]   (Gungbe, Aboh 2004)
          'I   heard that snake the **Top** Kofi killed it'
          'I think that, as for this book, you should read it tomorrow.'

     b.   Un sè   [do [dan lo **wɛ** [Kofi hu   _]]]   (Gungbe, Aboh 2004)
          'I   heard that snake the **Foc** Kofi killed '
          'I think that THIS BOOK you should read not the one by Suru.'

Under uniformity guidelines, it is natural to make the hypothesis that all languages involve a system of criterial heads, and what varies is their morphological realization, a low-level spell-out parameter. This argumentation of course presupposes that Top and Foc markers in cases like (2) indeed are part of the clausal spine, and are not postpositions or case-like endings attached to the moved phrases (see Rizzi 2013 for possible kinds of relevant evidence). Moreover, uniformity assumptions typically have to be weakened, as the analysis proceeds and gets refined, under the pressure of empirical evidence: obviously, natural languages are not completely uniform, so that non-trivial parameters of variation must be introduced; nevertheless starting from assumptions of uniformity is the necessary precondition for doing comparative syntax.[1]

At the LF interface, the criterial head guides the interpretation by triggering particular interpretive routines which capitalize on the transparent syntactic representations, e.g., for topics,

(4)   [   ] Top [   ]
        "Topic"   "Comment"

Or, informally, "interpret the specifier of Top as the topic, a contextually salient referent, and the complement of Top as a comment that is made about the topic". Interpretive routines at the semantic-pragmatic interface define the conditions for appropriate use in discourse. These conditions can be studied by setting up mini-discourse contexts, and testing the appropriateness of the criterial configurations in such discourse fragments. Consider, for instance, the following discourse fragment in Italian, illustrating licensing conditions on topics:

(5) a.   I libri che hai comprato ieri mi sembrano adatti per i regali di Natale…
          'The books you bought yesterday seem appropriate as Christmas
          present…'.
     b.   In effetti, il libro di Gianni, lo vorrei regalare a Maria.
          'In fact, Gianni's book, I would like to give to Maria'.

---

1. An extraordinary opportunity for developing such cross-linguistic arguments is offered by domains that show an important range of microvariation in historically related system (Kayne 2005), such as the dialects of Romance and Germanic, e.g., in the spell-out of heads of the high IP field and of the complementizer system: Poletto (2000), Manzini & Savoia (2005).

i.e., a left-peripheral topic selects a referent drawn from a presupposed set: in the case of the discourse fragment reported in (5), speaker A introduces a set of books, and interlocutor B selects a member of this set as a topic and makes a comment about it.

The Top head also guides contour assignment rules in the path to the PF interface, along the lines, e.g., if the system presented in Bocci (2013) for Italian.

Analogous considerations hold for left-peripheral focus. In Romance, not all kinds of foci can use left peripheral slots. For instance, simple new information focus (e.g., expressing the value of the variable in the answer to a wh question: Belletti 2001, 2004), uses an IP internal position in standard Italian and many other Romance varieties (see also Cruschina 2012 for elements of variation on this point). On the contrary, one kind of focus which can be expressed in the left periphery is corrective focus, which singles out a piece of information attributed to the interlocutor (typically expressed by the interlocutor in the immediate discourse context) and corrects it, as in the following discourse fragment, taken from Bianchi, Bocci and Cruschina (2014):

(6)   A: Se ho capito bene, sono andati alle isole Vergini.
              'If I understood correctly, they went to the Virgin Islands.'
        B: Ti sbagli! ALLE MALDIVE sono andati in viaggio di nozze!
              'You are wrong! TO THE MALDIVES they went on honeymoon!'

The interpretation here is determined by an interpretive routine like the following:

(7)   [   ] Foc_{contr} [   ]
       "Focus_{contr}"   "Presupposition"

i.e., against a presupposed background (they went somewhere on honeymoon), interlocutor B corrects interlocutor A's statement on the destination of their trip. The left peripheral position expressing the correction bears high prominence, whereas the presupposed part bears a flat intonational contour, as is shown in the following figure based on the experimental work presented in Bianchi. Bocci & Cruschina (2014), and discussed in Rizzi & Bocci (2016):

(8)   (Corrective) Focus – Presupposition (from Bianchi, Bocci & Cruschina 2014)
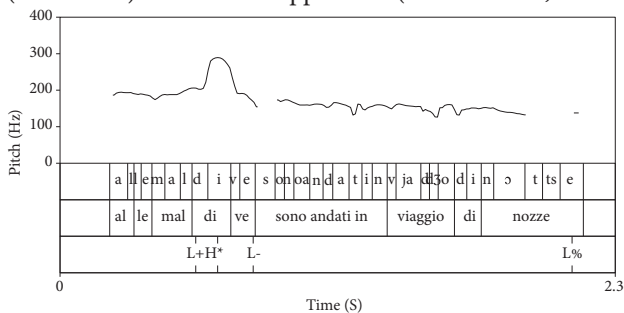
Bianchi, Bocci and Cruschina (2014) also show that Italian can use the left peripheral slot to express another kind of focus, mirative focus, through which the speaker expresses an information that is surprising and unexpected given his/her previous set of assumptions. Such a background licensing mirative focus can be typically expressed in the immediately preceding sentence in a discourse fragment like the following:

(9)   …E io che credevo che fossero dei poveracci! Figurati un po'…
    '…and I believed they were poor people! Can you imagine….
    … UNA FERRARI si sono comprati!
    '… A FERRARI bought for themselves!'

The contour of mirative focus is distinct from the one assigned to corrective focus, Bianchi et al., cit., argue, but the contour assigned to the presupposition of mirative focus shows the same flattening as the presupposition with corrective focus.

3.   There can be multiple topics, but LP focus is unique: The role of LF interface principles

Many languages permit a proliferation of topics, e.g. Italian (and Romance in general), and Abidji, spoken in Ivory Coast (Hager-Mboua 2014):

(10) A Maria, domani, il tuo libro, glielo devi dare al più presto
    'To Maria, tomorrow, your book, you it-to him should give as soon as possible'.

(11) kòfí **ɛ́kɛ́** òkókò ɛ́   **ɛ́kɛ́** è pìpjé nì.
    Kofi TOP banana DEF TOP ASPpeel RES PRON
    'Kofi, the banana, he peeled it.'

The Abidji case is particularly interesting in this connection, as the distinct topics all occur with the overt topic head *ɛ́kɛ́*, which strongly suggests that the proliferation of topics may be analyzed as a recursion of TopP (rather than as multiple specifiers attached to a unique Top head, or through multiple adjunctions to some left-peripheral constituent).

    Contrary to topic, multiple left-peripheral focus seems to be systematically proscribed. The uniqueness of LP focus has been observed for Italian (Rizzi 1997), English, Armenian (Giorgi & Haroutyunian, talk, University of Siena, 2016), Hungarian (Puskás 2000), Hebrew (Shlonsky 2014), Jamaican (Durrlemann 2008), Aboh (2004), Abidji (Hager – MBoua 2014). In the last three languages

the uniqueness is easy to determine, as these languages have overt focus markers, which can occur only once:[2]

(12) Jamaican: * A di bami   a di pikni im gi
                    The bammy the child he give   (Durrleman 2008: 75)

(13) Gungbe: * wémà lɔ **wὲ** Sɛna **wὲ** zé
                    THE BOOK   SENA   took   (Aboh 2004)

(14) Abiji: * òkókò$_i$ ɛ́   **bɛ́**   kòfí$_j$ **bɛ́** –   pìpjé –
                banana$_i$ Def. Foc Kofi$_j$ Foc – PEEL.RES –
            « THE BANANA, KOFI – peeled –_ »   (Hager-Mboua 2014)

In languages with null Foc, more care is needed to test the property, in order to distinguish LP focus from, e.g., contrastive topic, a notion that is sometimes confused with focus.

Consider the following discourse fragment in Italian, in which a multiple corrective focus could arise, but never is well formed. Speaker A produces (15)A; his interlocutor may want to correct two elements in this statement: perhaps it was Gianni, not Piero who won an important competition; moreover the competition was the world championship, not the Olympics. Nevertheless, speaker B cannot correct both elements in a single clause with two corrective foci: (15)B is ill-formed. The only possibility is to split his corrective statement into two sentences, each of which has a unique corrective focus, as in (15)B′:

(15) A. So che quest'anno Piero ha vinto le olimpiadi…
        'I know that this year Piero won the Olympics…'
     B. *Ti sbagli: quest'anno, GIANNI, I MONDIALI ha vinto, non Piero, le olimpiadi
        'You are wrong: this year, GIANNI, THE WORLD CHAMPIONSHIP won, not Piero, the Olympics.
     B′: Ti sbagli: quest'anno, GIANNI ha vinto una competizione importante, non Piero; e poi, I MONDIALI, ha vinto, non le olimpiadi
        'You are wrong: this year GIANNI won an important competition, not Piero; moreover, THE WORLD CHAMPIONSHIP he won, not the Olympics.

Mirative LP focus also is unique. In (16)A the speakers expresses a situation that was unexpected given his previous beliefs: that a member of his family would buy an expensive car. In continuation (16)B, mirativity is associated with the subject,

---

2. In Jamaican the focus marker *a* precedes the focalized element, rather than following it, a state of affairs that Durrleman (2008) analyzes as involving head movement of *a* from Foc to an immediately adjacent higher head position.

the unexpected agent of a contextually given predicate, buying an expensive car; in (16)B′ mirativity is associated with the object, the expensive car, in a context in which the topic is a member of my family, and the fact that he would buy a car is presupposed. Both choices are legitimate continuations of (16)A, which shows that there is some leeway, in a given context, in partitioning the structure between what is surprising and what is presupposed; but what is excluded is a double mirative focus, as in (16)B″.

(16) A. Non avrei mai pensato che qualcuno nella mia famiglia avrebbe speso tanti soldi per una macchina…
          'I would never had thought that anybody in my family would spend so much money for a car…'
     B.   E invece, figurati un po', proprio MIO FRATELLO ha comprato una Ferrari.
          'And then, could you believe it, precisely MY BROTHER bought a Ferrari'.
     B′   E invece, figurati un po', mio fratello UNA FERRARI ha comprato.
          'And then could you believe it, my brother A FERRARI bought'.
     B″ * E invece, figurati un po', proprio MIO FRATELLO UNA FERRARI ha comprato.
          'Ad then could you believe it, precisely MY BROTHER A FERRARI bought'.

I proposed in Rizzi (1997) that Uniqueness of LP focus follows from the interpretive procedure associated to the structure, repeated in (17) for convenience: if a FocP was recursively embedded as the complement of a higher Foc, we would have that the complement of a higher Foc (underscored in (18)), a presupposition according to (17), contains a focus position, an inconsistent interpretive property.

(17) [   ] Foc [   ]
          "Focus"   "Presupposition"

(18) *[A MARIA] **Foc1** [[IL TUO LIBRO] **Foc2** [devi dare]]]
          'To MARIA   YOUR BOOK   you should give

On the other hand, no interpretive requirement blocks recursion of topic – comment structures: the only requirement on the comment may be that it contains new information, but this is consistent with a (reiterated) topic – comment structure.

(19) [A Maria] **Top1** [[il tuo libro] **Top2** [glielo devi dare]]]
          'To Maria   your book   you it-to-him should give

So, recursion of FocP is generally excluded by the interpretive procedure triggered by the Foc head at LF, whereas no general interpretive incompatibility arises in the

case of TopP recursion: other factors may require uniqueness of Topics in some languages (see Rizzi 2013), but there is no general ban against topic recursion, and in fact multiple topics are possible in many languages.

4.   Principles and parameters: A role for PF interface?

In Italian, the restriction to a single LP focus holds not only in simple clauses, but also in complex sentences (in fact, restrictions on co-occurrence of foci may hold more generally, but in this paper I will only look at the cases of LP focus). So, for instance, we can have a corrective focus in a main clause (as in (20)a), or in an embedded clause (as in (20)b), but not simultaneously in the main and the embedded clause (as in (20)c):

(20) a. A GIANNI ho detto – che dovremmo leggere il tuo libro, non a Piero
          'TO GIANNI I said that we should read your book, not to Piero'.
      b. Gli ho detto che IL TUO LIBRO dovremmo leggere –, non quello di Franco
          'I said to him that YOUR BOOK we should read, not Franco's".
      c. *A GIANNI ho detto – che ILTUO LIBRO dovremmo leggere –, non a Piero, quello di Franco
          'TO GIANNI I said that YOUR BOOK we should read, not to Piero, Franco's'.

In Rizzi (1997, fn. 15) I proposed that this restriction also follows from interpretive routine (17): the LP focus in the lower clause in (20)c is included in the presupposition of the higher focus (if the presupposition is the whole c-domain of the Foc head), therefore a clash arises also when the two foci are not in the same LP.

But comparative considerations become relevant here. If two LP foci in the same clause are systematically banned across languages, the co-occurrence of LP foci in different clauses of the same complex sentence is clearly permitted in some languages. A case in point is Gungbe (Aboh 2004). A *wɛ* marked phrase can appear in a main clause and in an embedded clause in the same complex sentence:

(21) a. Sena **wὲ** _ sὲ   dɔ Remi **wὲ** _ zé   hi lɔ
          Sena Foc   hear-Perf that Remi Foc – take-Perf knife + def
          'SENA heard that REMI took the knife'.
      b. Sena **wὲ** _ sὲ   dɔ hi lɔ   **wὲ** Remi zé   _
          Sena Foc   hear-Perf that knife + def Foc Remi take + perf
          'SENA heard that Remi took THE KNIFE'.

So, some kind of parametrization must be at play here. One possibility that immediately comes to mind is that the parametrization could involve the calculation of the presupposition:

(22) the presupposition associated to Foc extends
    (i) to the simple clause c-commanded by Foc (Gungbe)
    (ii) to the whole complex sentence c-commanded by Foc (Italian)

If this assumption is made, the interpretive clash between being focal and being presupposed would continue to arise for complex clauses in Italian like (20)c, but not in the Gungbe examples (21), where the calculation of the presupposition would start anew in the embedded clause.

Nevertheless, this approach seems to raise serious learnability issues: how would the language learner come to determine the language-specific parametric value of (22)? A more promising avenue would be to try to connect the observed difference to some salient difference between the two languages, immediately accessible to the language learner.

A clear difference between the two languages exists in the properties of the PF interface of the focus constructions. We have seen that in Italian a very special intonational contour is assigned to structures involving LP focus, as shown by (8). The PF prosody is characterized, among other properties, by the flattening of the contour of the presupposition. This may be inconsistent with the assignment of another focal prominence in the flattened part, and this sole PF inconsistency may be what goes wrong in (20)c. This is not an obvious line of analysis (see Bocci 2013 for discussion), but it seems to me to be intuitively plausible enough to be pursued.

Contrary to Italian, no special contour assignment is operative in Gungbe. Aboh writes:

(23) "…no stress mechanism arises in the Gungbe focus strategy. Focusing is realized only through movement of the focused element to the left-adjacent position to **wɛ**…" (Aboh 2004: 238)

How can we connect the contrast illustrated by (20)–(21) to the independent difference between Italian and Gungbe emerging from (23)? One natural possibility is to assume that the calculation of presupposition at the LF interface, contrary to approach (22), is not parametrized. Suppose that the calculation of the presupposition is always local, as in (22)i. In fact this is the expected state of affairs if the process is phase-based: the system would only 'see' the content of the local clausal phase, without penetrating more embedded clausal phases.[3]

---

3. If phase nodes are CP and vP, or, using finer distinctions, ForceP and vP, the system should be able to "see" ForceP + vP configurations. This would be consistent with, e.g., the mechanism

If the calculation of the presupposition can see the whole content of simple clauses, the double occurrence of a LP focal position in a single clause is systematically excluded by the interpretive clash that would arise at LF. So, the general state of affairs found uniformly in Gungbe, Italian and the other languages mentioned in Section 3 can be captured as a general property of natural language.

What about the variation observed in (20)–(21)? It could not be captured by the mechanism at the LF interface, now assumed to be invariant, but it follows from the independently observed difference at the PF interface: Italian has contour assignment rules that are incompatible with another focus in the flattened string following a focus position, no matter how distant the second focus is from the first,[4] whence the ill-formedness of (20)c. On the contrary, in Gungbe, as no special contour assignment rule applies, nothing excludes the co-occurrence of a focus in the main clause and another non-local focus in an embedded clause, as in (22)a-b.

In conclusion, by assuming an invariant syntax and an invariant mapping to LF, the system captures the general property of the uniqueness of left-peripheral focus in each simple clause. By assuming variation where it can be naturally expected and is widely attested, in the mapping to PF, the system captures the observed difference between languages allowing a single LP focus per complex sentence, like Italian, and languages allowing LP focus positions in different clauses of the same complex sentence, like Gungbe.

Acknowledgments

---

postulated in Chomsky (2001), or with some mechanism of 'phase sliding' (Gallego 2010) making the content of a simple clause globally accessible to the computation.

4. An anonymous reviewer observes that the ill-formedness of examples like (20)c is significantly weakened if the higher focus is expressed through the cleft construction:

i. ? E' a Gianni che ho detto che IL TUO LIBRO dovremo leggere, non a Piero, quello di Franco
    'It is to Gianni that I said that YOUR BOOK we should read, not to Peter, Franco's'

The reviewer observes that this improvement is expected under the PF interface approach to (20)c because the cleft construction does not seem to require the contour flattening that is enforced with focus preposing to the left periphery. Importantly, the contrast (20)c-(i) is captured by the PF account, but it would not be expected under the LF parametrization account in (22).

Leonardo Savoia, who was then *perfezionando* at the *Scuola*. Our discussions were nourished by the breadth of his interests, spanning from syntax to phonology, semantics, dialectology, and more. I hope Leonardo will find in these pages some echoes from these early exchanges and later interactions, and will enjoy reading an attempt to combine a formal syntactic analysis with the study of the interfaces of syntax with sound and meaning.

References

Aboh, Enoch Olade. 2004. *The Morphosyntax of Complement-head Sequences: Clause Structure and Word Order Patterns in Kwa*. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780195159905.001.0001

Belletti, Adriana. 2001. Inversion as focalization. In *Subject Inversion in Romance and the Theory of Universal Grammar* , Aafke C.J. Hulk & Jean-Yves Pollock (eds) 60–90. Oxford: OUP.

Belletti, Adriana. 2004. Aspects of the low IP area. In *The Structure of CP and IP. The Cartography of Syntactic Structures*, Vol. 2, Luigi Rizzi (ed.) 16–51. Oxford: OUP.

Bianchi, Valentina, Giuliano Bocci & Silvio, Cruschina. 2014. Focus and its implicatures. In *Romance Languages and Linguistic Theory: selected papers from Going Romance 2013*,[Romance Languages and Linguistic Theory 8] Enoch Aboh, Jeannette Schaefer & Petra Sleeman (eds) 1–20. Amsterdam: John Benjamins.

Bocci, Giuliano. 2013. *The Syntax–Prosody Interface: A Cartographic Perspective with Evidence from Italian*.[Linguistik Aktuell/Linguistics Today 204] Amsterdam: John Benjamins.
https://doi.org/10.1075/la.204

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

Cinque, Guglielmo & Rizzi, Luigi. 2010. The cartography of syntactic structures. In *The Oxford Handbook of Linguistic Analysis*, Bernd Heine & Heiko Narrog (eds). Oxford: OUP.

Cruschina, Silvio. 2012. *Discourse-related Features and Functional Projections*. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199759613.001.0001

Durrleman, Stephanie. 2008. *The Syntax of Jamaican Creole*.[Linguistik Aktuell/Linguistics Today 127] Amsterdam: John Benjamins. https://doi.org/10.1075/la.127

Gallego, Ángel J. 2010. *Phase Theory* [Linguistik Aktuell/Linguistics Today 152] Amsterdam: John Benjamins. https://doi.org/10.1075/la.152

Hager M'Boua, Clarisse. 2014. Structure de la phrase en Abidji. PhD dissertation, University of Geneva.

Kayne, Richard. 2005. *Parameters and Universals*. Oxford: OUP.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci*. Alessandria: Edizioni dell'Orso.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. Oxford: OUP.

Puskás, Genoveva. 2000. *Word Order in Hungarian: The Syntax of Ā-positions* [Linguistik Aktuell/Linguistics Today 33] Amsterdam: John Benjamins. https://doi.org/10.1075/la.33

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar: A Handbook of Generative Syntax*, Liliane Haegeman (ed.) 281–337. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25(1): 197–226.

Rizzi, Luigi. 2017. The left periphery: Cartography, freezing, labeling. In *Studies on Syntactic Cartography*, Fuzhen Si (ed.). Beijing.

Rizzi, Luigi & Cinque, Guglielmo. 2016. Functional categories and syntactic theory. *Annual Review of Linguistics* 2: 139–163. https://doi.org/10.1146/annurev-linguistics-011415-040827

Rizzi, Luigi & Bocci, Giuliano. 2015. The left periphery of the clause – Primarily illustrated for Italian. In *The Blackwell Companion to Syntax*, 2nd edn, Martin Everaert & Henk van Riemsdijk (eds). Malden MA: Wiley-Blackwell.

Shlonsky, Ur. 2014. Topicalization and focalization: A preliminary exploration of the Hebrew left periphery. In Peripheries, Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds), 327–341. Tokyo: H. Syobo.

CHAPTER 11

Italian *faire*-infinitives

The special case of *volere*

Norma Schifano and Michelle Sheehan
University of Cambridge / Anglia Ruskin University

The first aim of this squib is to show that, under the correct syntactic configuration, *volere* 'want' can be embedded in Italian *faire*-infinitives (FI), contrary to previous claims. Secondly, we show that want-FIs exhibit peculiar properties: (i) they disallow a full DP causee; (ii) they permit intermediate cliticisation onto *volere*; (iii) they allow optional splitting of clitic clusters; and (iv) they marginally permit an accusative causee in transitive contexts where the object is a clitic. We attribute these effects to defective intervention which bans a full DP causee and requires the creation of a biclausal ECM construction, where accusative is exceptionally licensed, and where the selection of a silent OBTAIN by 'want' creates an additional clitic position.

Keywords: Italian, *faire*-infinitives, want, clitics, ECM constructions

1. Introduction

The aim of this squib is to bring to light some curious facts regarding selection and case in the Italian *faire-infinitive* causative construction (FI) which have not, to our knowledge, been hitherto discussed. More specifically, we will show that there are contexts in which causative *fare* can embed the volitional restructuring verb *volere* 'want', contrary to the predictions of Cinque's (2006) hierarchy of restructuring predicates, and more in line with biclausal analyses of causatives (see Brandi & Savoia 1990; Guasti 1993; Manzini & Savoia 2007 amongst others). Such examples have interesting syntactic properties, moreover, for which we sketch an analysis. We begin with some background on FIs in Section 2. Section 3 presents the interesting syntactic properties of want-FIs. Section 4 outlines an analysis of the ban on full DP causees in terms of defective intervention. Section 5 turns to the possibility of ECM in contexts where both the object and the causee are cliticised. Section 6 addresses the splitting of the clitic cluster. Finally, Section 7 concludes.

2.   Embedding *volere* in Italian *faire*-infinitives

Based on extensive cross-linguistic evidence, Cinque (2003, 2004, 2006) proposes that functional predicates occupy dedicated positions within a strictly ordered hierarchy, as partially represented below (based on Ledgeway forthcoming):

(1)   [Mod$_{Epistemic/Alethic}$ *dovere/potere* … [Mod$_{Volition}$ *volere* … [Mod$_{Obligation/Ability}$ *dovere/potere* [Mod$_{Permission}$ *potere* … [**Causative** *fare* [Asp$_{Inceptive}$ *cominciare* [Asp$_{Andative}$ *andare* [Asp$_{Completive}$ *finire* [$_{v\text{-}VP}$ V…

Italian causative verbs like *fare* 'make', Cinque (2006, 72ff) places in a relatively low position, following the observation that they can be passivized (2a) but cannot embed a passive (2b), suggesting that Voice°, which is placed very low in Cinque's hierarchy (Cinque 1999: 106), outscopes Causative° (see (1) and the smaller portion of the hierarchy in (3)):

(2)   a.   Gli   fu   fatto leggere   (Cinque 2006: 72)
          3SG.DAT= it.was made read.INF
          'He was made to read it'
      b.   *Farò   essere invitati tutti
           I.will.make be.INF invited all
           'I will make everybody be invited'

(3)   … Voice° > … Causative° > Asp$_{inceptive(II)}$ /
      (Asp$_{continuative(II)}$) > Andative° > Asp$_{completive(II)}$ (based on Cinque 2006: 76)

A strong prediction of this approach is that restructuring verbs higher than Voice should not be able to be embedded under a causative verb. This is certainly the case for *avere* have/*essere* be/*venire* come (Burzio 1986; Guasti 1993, 1996; Folli & Harley 2007):

(4)   *Farò   aver riparato la   macchina a Gianni, per domani
       I.will.make have repaired the car   to Gianni by tomorrow
                                                                   (Guasti 1996: 298)
       Intended: 'I will make Gianni have repaired the car by tomorrow'

The expectation also seems to be borne out by the data reported by Cinque (2006, 72ff) (see also Burzio 1981: 587). Compare (5), where *volere*, *dovere*, and *potere* apparently cannot be embedded under *fare* with (6), where aspectual *cominciare* 'begin', *andare* 'go' and *finire* 'finish', which are lower in (1), can be:

(5)   a.   *La   feci   **voler**   leggere a tutti
            3SG.ACC= I.made want.INF read.INF to everyone
            'I made everybody want to read it'   (Cinque 2006: 72–73)

b. *Lo   faranno   **dover**   ammettere anche a Gianni
3SG.ACC= they.will.make must.INF admit.INF also   to G.
'They will make Gianni too have to admit it'

c. *Lo   farò   **poter**   leggere a tutti
3SG.ACC= I.will.make can.INF read.INF to everyone
'I will make everybody be able to read it'

(6) a. Gliela   fecero   **cominciare** a costruire
3SG.DAT=3SG.ACC= they.made begin.INF   to build.INF
'They had him begin to build it'   (Cinque 2006: 74)

b. Ce   lo   fecero   **andare** a prendere subito
1PL.DAT= 3SG.ACC= they.made go.INF to fetch.INF immediately
'They made us go and fetch it immediately'

c. La   fecero   **finire**   di costruire a Gianni
3SG.ACC= they.made finish.INF of build.INF to G.
'They had Gianni finish building it'

It is possible, however, to construct different examples with *volere*, which are far more acceptable:[1]

(7) a. È la cosa che ti   fa   **voler**   scrivere la storia
it.is the thing that 2SG.DAT= makes want.INF write.INF the history
'It is the thing that makes you want to write history'

b. Sono sicuro di non sapere,   Herr Issyvoo, cosa le   fa
I.am sure   of not know.INF H.   I.   what 3SG.DAT= makes
voler   lasciare Berlino
want.INF leave.INF Berlin
'I'm sure I don't now, Herr Issyvoo, what makes you want to leave Berlin'

c. E   quindi, sì, questo mi   fa   **voler**   conoscere il   mio
and thus   yes this   1SG.DAT= makes want.INF know.INF the my
vero padre, a volte
real father at times
'And thus, yes, this makes me want to meet my real father, sometimes'

One crucial difference between Cinque's ungrammatical example in (5a) and the grammatical forms in (7) is the syntactic expression of the causee. While in (5a) the causee is realised as a full DP (cf. *a tutti*), in (7) it is cliticised onto *fare*. In

---

1. Judgments vary considerably among speakers. The examples in (7) were retrieved from the Internet by the authors (1/09/2017). (7a)–(7b) were judged fully grammatical by 5/5 native speakers with who we carried out a preliminary questionnaire, while (7c) was judged fully grammatical by 3 of them and partially grammatical by 2. Note that the examples in (7) are also problematic for Folli and Harley's (2007) influential analysis of FIs as the causees are not agents.

what follows, we discuss this property of want-FIs and its repercussions, comparing them with prototypical FIs.

3.   Want-FIs: Syntactic properties

In a prototypical transitive FI like (8a), the direct object, if cliticised, must climb to *fare* (8b):

(8) a. Il parroco fa   costruire un tavolo da ping pong al   falegname
          the priest   makes build.INF a   table   of ping pong to.the carpenter
          'The priest makes the carpenter build a ping pong table'
     b. Il parroco **lo**   fa   costruire al   falegname
          the priest   3SG.ACC= makes build.INF to.the carpenter
          'The priest makes the carpenter build it'

Similarly, the causee (S_2) if cliticised, must also climb to *fare*, being accusative with intransitive and dative with transitive caused events, as in (9):

(9) Il parroco **gli / *lo**   fa   costruire un tavolo da ping pong
       the priest   3SG.DAT=/ACC= makes build.INF a   table   of ping pong
       'The priest makes him build a ping pong table'

If both the direct object and S_2 are clitics, both climb onto *fare* (10a). If only one of the two undergoes climbing, the sentence becomes ungrammatical, regardless of whether the stranded clitic is the accusative (10b) or the dative (10c), mimicking the behaviour of clitics in restructuring contexts:

(10) a. Il parroco **glielo**   fa   costruire
            the priest   3SG.DAT=3SG.ACC= makes build.INF
       b. *Il parroco **gli**   fa   costruir**lo**
            the priest   3SG.DAT= makes build.INF=3SG.ACC
       c. *Il parroco **lo**   fa   costruir**gli**
            the priest   3SG.ACC= makes build.INF=3SG.DAT
            'The priest makes him build it'

If the same diagnostics are applied to what we will call 'want-FIs', we obtain different results. First, want-FIs with a non-pronominal S_2 are highly degraded/ungrammatical, whether the embedded object is a full DP (11a), or a clitic, and regardless of the latter's placement (see Cinque's 5a and 11b):

(11) a. ?*Non so   cosa fa   voler   lasciare Berlino al   signor Herr
           not I.know what makes want.INF leave.INF Berlin   to.the mister H.
           'I don't know what makes Mr Herr want to leave Berlin'

   b. *Non so   cosa **la**   fa   voler(**la**)
        not I.know what 3FSG.ACC= makes want.INF(3FSG.ACC=)
        lasciar(**la**)   al   signor Herr
        leave.INF(3FSG.ACC=) to.the mister H.
        'I don't know what makes Mr Herr to want to leave it/her'

Where S$_2$ is a pronominal clitic, however, the sentence becomes grammatical, whether the clitic climbs onto *fare* (12a) or, more markedly, to *volere* (12b) (vs 12c).[2] Note that the S$_2$ clitic, as expected, obligatorily surfaces in the dative, where the object is a full DP, as this is a transitive context:

(12) a.   Non so   cosa **gli** (/*lo**)   fa   voler   lasciare
             not I.know what 3SG.DAT= (/3SG.ACC=) makes want.INF leave.INF
             Berlino
             Berlin

   b. (?)Non so   cosa fa   voler**gli**(/*lo**)   lasciare
        not I.know what makes want.INF=3SG.DAT(/3SG.ACC) leave.INF
        Berlino
        Berlin

   c. *Non so   cosa fa   voler   lasciar**gli/lo**
        not I.know what makes want.INF leave.INF=3SG.DAT/=3SG.ACC
        Berlino
        Berlin
        'I don't know what makes him want to leave Berlin'

Finally, if both the direct object and S$_2$ are cliticized, they can climb as a cluster onto *fare* (13a), as with transitive FIs (see (10a) above), or, less preferably, to *volere* (13b) (vs (13c)):

(13) a.   Sono sicuro di non sapere   cosa **gliela**   fa   voler
             I.am sure   of not know.INF what 3SG.DAT=3SG.ACC= makes want.INF
             lasciare
             leave.INF

   b. (?)Sono sicuro di non sapere   cosa fa   voler**gliela**
        I.am sure   of not know.INF what makes want.INF.=3SG.DAT=3SG.ACC=
        lasciare
        leave.INF

---

2. Further investigation is required to assess the acceptability of examples like (12b).

    c. * Sono sicuro di non sapere   cosa fa   voler
        I.am sure   of not know.INF what makes want.INF
        lasciar**gliela**
        leave.INF.=3SG.DAT=3SG.ACC
        'I'm sure I don't know what makes him want to leave it'

Interestingly, moreover, for some speakers, want-FIs also (marginally) allow a split option whereby only S$_2$ climbs to *fare*, leaving the object clitic attached to *volere* (14a) (vs 10b). Even more strikingly, in such contexts, S$_2$ can exceptionally surface as accusative (14b) (vs 9):[3]

  (14) a. $^{?}$ Sono sicuro di non sapere   cosa **gli**   fa   voler**la**
             I.am sure   of not know.INF what 3SG.DAT= makes want.INF=3SG.ACC
             lasciare (/*?lasciar**la**)
             leave.INF /leave=3SG.ACC
      b. $^{?}$ Sono sicuro di non sapere   cosa **lo**   fa   voler**la**
             I.am sure   of not know.INF what 3SG.ACC= makes want.INF=3SG.ACC
             lasciare (/*?lasciar**la**)
             leaveINF /leave=3SG.ACC
             'I'm sure I don't know what makes him want to leave it'

Climbing of the direct object, leaving accusative/dative S$_2$ stranded, is banned for all speakers:

  (15) *Sono sicuro di non sapere   cosa **la**   fa
       I.am sure   of not know.INF what 3SG.ACC= makes
       voler**gli**(/**lo**)   lasciar**gli**(/**lo**)
       want.INF=3SG.DAT(/=3SG.ACC) leave.INF=3SG.DAT(/=3SG.ACC)

The relevant patterns are summarised in Table 1:[4]

---

3. Our preliminary survey with 5 native speakers revealed a great deal of variation in the acceptability of (14a)–(14b). Dative marking (14a) was rated ungrammatical by 3 speakers and partly or fully grammatical by 2 speakers and accusative marking (14b) was rated ungrammatical by 3 speakers and grammatical by 2 speakers. Further investigation in required to assess this point and to clarify whether regional variation may play a role.

4. O = object; CC = clitic climbing.

Table 1. Cliticitisation in FIs and want-Fis

	clitic O		full O	
	full S_2	clitic S_2	full S_2	clitic S_2
FI	✔ (8b)	✔ both CC (10a) * 1 CC (10b)–(10c)	✔(8a)	✔ S_2.DAT (9)
want-FI	* (11b)	✔ both CC (13a) (?)both on *volere* (13b) ? only S_2.DAT CC (14a) ? only S_2.ACC CC(14b)	?* (11a)	✔S_2.DAT on *fare* (12a) (?) S_2.DAT on *volere* (12b)

These patterns raise four interesting questions: (i) why is it the case that the causee of *volere* can only be expressed as a clitic and never as a full DP (cf. (11) vs (12))?; (ii) why do want-FIs permit intermediate cliticisation onto *volere*? (12b, 13b)?; (iii) why is it the case that want-FIs allow split clitic climbing for some speakers (14)? and (iv) why is it the case that, with a split clitic cluster, S_2 can be accusative (14b) rather than dative (14a) for some speakers? In what follows, we suggest potential answers to some of these questions.

4. Clitic transitive causee: defective intervention

The fact that transitive causees can only occur as clitics in want-FIs (cf. i) might result from defective intervention.[5] Abstracting away from some details, let us assume that in FIs, the object can usually be probed by *fare* and receive Case from it either through leapfrogging (16a) (McGinnis 1998; Sheehan, in progress) or smuggling (16b) (Collins 2005; Belletti & Rizzi 2012).[6] In want-FI contexts, however, this facilitating movement does not take place for some reason, and so the dative causee acts as a defective intervener for the Agree relation between *fare* and the embedded object (16c):

---

5. Although Bruening (2014) raises certain objections against the data originally used to argue in favour of defective intervention, Marchis Moreno & Petersen (2016, 2017) have defended its existence, albeit in a slightly different 'linear' form. For our purposes, what is crucial is that full DPs pose a problem for probing whereas clitics do not. We leave a full discussion of these matters to future research.

6. As an anonymous reviewer notes, this view departs from earlier approaches in which the object received Case from the entire verbal complex (complex predicate) (Guasti 1993), or by virtue of being in a single Case-domain with another argument (Folli & Harley 2007). Under such approaches, there should be no defective intervention effect and the ban on DPs causees would need to be explained in some other way.

(16) a.   [*fare*$_{[uPHI]}$ [DP$_{[uCase]i}$ **DP$_{DAT}$** *lasciare* [V t$_i$]]] (leapfrogging)
    b.   [*fare*$_{[uPHI]}$ [[V DP$_{[uCase]}$]$_i$ **DP$_{DAT}$** *lasciare* t$_i$]] (smuggling)
    c.   [*fare*$_{[uPHI]}$ [**DP$_{DAT}$** *volere lasciare* [V DP$_{[uCase]}$]]] (defective intervention)

The same problem arises in raising contexts and, in such cases, cliticisation of the dative removes the dative as a c-commander, bleeding the effect (Rizzi 1986; McGinnis 1998, 2000):

(17) a. $^{??}$Gianni sembra a Piero [t fare   il suo dovere]
       G.   seems   to P.   do.INF the his duty

                                        (McGinnis 1998: 92)

    b.   Gianni non **gli**   sembra [t fare   il suo dovere]
       G.   not 3SG.DAT= seems   do.INF the his duty
       'Gianni does not seem to Piero / him to do his duty'
                          (McGinnis 2000, citing Rizzi 1986)

Arguably the same effect occurs with want-FIs: full DP causees block Agree between *fare* and the embedded object, whereas clitic causees do not. This explains why full *dative* DP causees are blocked, but it remains to be shown why full *accusative* DP causees should also be banned in the ECM construction (e.g. **Non so cosa fa Herr (voler) lasciare Berlino* 'I don't know what makes Herr (want to) leave Berlin'). We consider this in the next section.

5.   Accusative transitive causee: ECM construction

Regarding (iii), the exceptional accusative marking of clitic S$_2$, we claim that this configuration instantiates a biclausal ECM construction. This may seem an odd state of affairs, given that Italian, on a par with French and unlike some Spanish varieties, notably does not allow ECM complements to *fare* (e.g. Guasti 1993; Folli & Harley 2007: 221, fn 23; Sheehan 2016: 985):

(18) a. *Ho   fatto Herr lasciare Berlino
       I.have made H.   leave.INF Berlin
    b. *L'ho   fatto lasciare Berlino
       3SG.ACC=have made leave   Berlin
       'I made Herr / him leave Berlin'

Interestingly, ECM does occur, however, in another causative context in Italian, suggesting that an ECM-analysis for accusative S$_2$ in want-FIs may be on the right track. As widely observed in the literature, Italian is subject to the 'person-case constraint' (PCC), banning a 3rd person indirect clitic with 1st/2nd person di-

rect one (see Perlmutter 1971; Bonet 1991) (19a). This effect extends to causative constructions, as shown in (19b):[7]

(19) a. *Lui **mi** **gli** presenta
he 1SG.ACC= 3SG.DAT= introduces
'He introduces me to him'

b. *Questa notizia **gli** **ci** ha fatto chiamare
this news 3SG.DAT= 1PL.ACC= has made call.INF
'This news made him call us'

A curious difference between PCC in ditransitives and causatives is that in causatives even a full indirect object DP is banned with 1st/2nd person direct object clitic, as noted by Sheehan (to appear), unlike in ditransitive contexts, where only clitics are problematic (as in (19a)):

(20) a. *Maria **mi** fece picchiare a Carlo
M. 1SG.ACC= made beat.INF to C.

b. ***Mi** **gli** fece picchiare
1SG.ACC= 3SG.DAT= he.made beat.INF
Intended: 'M. made Carlo/him beat me'

One possible repair for (20b), for some speakers, is to make the causee accusative. In such cases, only S$_2$ climbs, suggesting that this is a biclausal ECM construction (21a). Climbing of the direct object, instead of S$_2$ (either dative or accusative), yields ungrammaticality (21b), as does climbing of both clitics:[8]

(21) a. **Lo** / ***gli** fece picchiar**mi**
3SG.ACC= 3SG.DAT= (s)he.made beat.INF=1SG.ACC
'(S)he made him beat me'

b. ***Mi** fece picchiar**gli/lo**
1SG.ACC= (s)he.made beat.INF=3SG.DAT/=3SG.ACC

c. ***Me** **lo** fece picchiare
3SG.ACC= 3SG.ACC= (s)he.made beat.INF
Intended: '(S)he made him beat me'[9]

---

7. The examples in (19) are adapted from Roberts (2016: 789–790).

8. Speakers' judgements vary also in this domain. 3 out 5 speakers judged the accusative in (21a) as completely grammatical, while 2 judged it ungrammatical. 5 out of 5 speakers judged the dative option ungrammatical. This is consistent with the claim that not all speakers have ECM available as a repair mechanism.

9. (21c) is ungrammatical in the intended meaning under which *me* is accusative and *lo* dative. Note that the switch from *mi* to *me* simply follows from the morphophonemic adjustment of /i/ to /e/ before sonorants to which Italian clitic clusters are subject (Vincent 1988: 291–292).

Want-FIs in PCC-contexts behave in exactly the same way: a full indirect object DP is banned (22a), only S_2 can climb (22b), and it is obligatorily case-marked accusative (22c), while non-PCC want-FIs allow climbing of both, or marginally climbing of dative/accusative S_2 only (see Table 1):[10]

(22)   a. * **Mi**   fece   voler   picchiare a Carlo
             1SG.ACC= (s)he.made want.INF beat.INF to Carlo

       b. * **Mi**   **gli**   fece   voler   picchiare
             1SG.ACC= 3SG.DAT= (s)he.made want.INF beat.INF

       c.   **Lo**   (/***gli**)   fece   voler**mi**   picchiare
             3SG.ACC= 3SG.DAT= (s)he.made want.INF=1SG.ACC beat.INF
             '(S)he made Carlo / him to want to beat me'

All patterns thus far are summarised in Table 2:

Table 2. Cliticisation in FIs and want-FIs: PCC and non-PCC contexts

	clitic O		full O	
	full S_2	clitic S_2	full S_2	clitic S_2
FI	✓(8b)	✓ both CC (10a) *1 CC (10b)-(10c)	✓(8a)	✓ S_2.DAT (9)
FI (PCC)	*(20a)	*both (20b) ✓ only S_2.ACC CC (21a)		
want-FI	*(11b)	✓ both CC (13a) (?)both on *volere* (13b) ? only S_2.DAT CC (14a) ? ony S_2.ACC CC (14b)	?*(11a)	✓S_2.DAT on *fare* (12a) (?) S_2.DAT on *volere* (12b)
want-FI (PCC)	*(22a)	*both (22b) ✓ only S_2.ACC CC (22c)		

It should be clear from Table 2 that both FI and want-FI are subject to the PCC in the same way and yet in other contexts behave very differently. This poses certain challenges for analyses of the PCC which we will not go into here. For our purposes here, these facts illustrate that ECM is available under causative *fare*, though it is highly restricted. It remains unclear why it is so restricted. In PCC contexts, it can be considered a repair strategy, but with want-FIs it does not have this function as accusative clitic causees alternate with datives. It also remains unexplained why ECM is *only* possible with embedded clitic objects and not full object DPs (cf. 12a).

---

10. 3 out of 5 speakers judged the accusative option in (22c) fully grammatical, 1 judged it partially grammatical and 1 ungrammatical. 5 out of 5 speakers judged the dative option ungrammatical.

A further remaining question, mentioned above, is the following: as Italian allows ECM constructions in want-FIs (cf. 14b), why is the case that the causee of *volere* cannot be expressed as a full DP in ECM configurations (cf. **Non so cosa fa Herr voler lasciare Berlino* 'I don't know what makes Herr (want to) leave Berlin')? This cannot be attributed to defective intervention (cf. § 4) as ECM constructions are biclausal, with the embedded object being licensed in the lower clause and S_2 in the higher clause, by *fare*. We leave these intriguing questions open here.

6.   Intermediate cliticisation: Silent OBTAIN

We conclude by addressing the two issues raised by want-FIs in relation to cliticisation, i.e. the fact that for some speakers they marginally allow intermediate cliticisation onto *volere* (cf. ii, 12b, 13b) and splitting of the clitic cluster (cf. iii, 14). Starting from intermediate cliticisation, this is particularly surprising given that: (i) *fare* usually forces clitic climbing (Kayne 1975; Rizzi 1976: 42; Burzio 1986: 260ff; Cardinaletti & Shlonsky 2004: 543; Ledgeway 2016: 224; Roberts 2016: 800; Sheehan 2016: 986); (ii) if more than one restructuring predicate is present (here *fare* and *volere*), the clitic must climb onto the highest (e.g. *lo devo riuscir(**lo) a fare* '3SG.ACC= I.must manage(=3SG.ACC) to do', Ledgeway 2016: 223). As Cardinaletti and Shlonsky (2004) note, descriptively speaking, prototypical functional verbs embedded in restructuring contexts do not allow intermediate climbing (unless restructuring stops at the intermediate position), because they are not associated with a clitic position. The only verbs to be associated with a clitic position are: (i) lexical verbs (cf. *voglio mangiarlo* 'I want to eat it'), and (ii) a special class of functional verbs, which they call 'quasi-functional', which includes causative, motion and perception verbs.[11] These verbs exceptionally introduce an extra clitic position, thus allowing intermediate climbing, as shown by the contrast below, where *andare* 'go' (23a) but not *potere* 'can' (23b) can host the clitic:

(23) a. Sarei   voluto poter andar**lo**   a trovare
          I.would.be wanted can.INF go.INF=3SG.ACC to meet.INF
      b. *Sarei voluto poter**lo** andare a trovare
          'I would have wanted to be able to go and visit him'

                                    (Cardinaletti & Shlonsky 2004: 541)

---

11. There is a further clitic position, in addition to the one associated with lexical and semifunctional verbs, which is a high position in the IP, where the clitic appears in contexts of climbing (e.g. **lo** *voglio mangiare* 'I want to eat it') (Cardinaletti & Shlonsky 2004).

We observe here that the same contrast applies to causatives: when *fare* combines with lower functional verbs, *cominciare* cannot host a clitic (24), whereas semi-functional *andare* can (25):

(13) a. *Il parroco fece cominciar**gli**   a costruire un tavolo
           the priest   made start.INF=3SG.DAT to build.INF a   table
           da ping pong
           of ping pong

      b. Il   parroco **gli**   fece   cominciare a costruire un tavolo
           the priest   3SG.DAT= made start.INF   to build.INF a   table
           da ping pong
           of ping pong
           'The priest made him start building a ping pong table'

(25) a. **Mi**   fecero andar**la**   a prendere subito
           1SG.DAT= made GO.INF=3SG.ACC to fetch.INF immediately

      b. **Me**   **la**   fecero andare a prendere subito
           1SG.DAT= 3SG.ACC= made go.INF to fetch.INF immediately
           'They made me go to fetch her immediately'

As intermediate clitic climbing is also allowed when *fare* combines with *volere* (12b, 13b), we must conclude that *volere* too is associated with an extra clitic position. This is unexpected, as *volere* is neither a lexical verb, nor a verb falling into the quasi-functional category of Cardinaletti & Shlonsky's (2004) classification. However, as Cinque (2006: 17) notes, *volere* is unique with respect to other functional verbs in other ways too. For example, it is alone in allowing two uses of adverbs like *già* 'already' and *sempre* 'always' with an infinitive complement (such examples are disallowed with *dovere/cominciare*):[12]

(26) Maria vorrebbe **già**   averlo   **già**   lasciato
      M.   wants.COND already have=3SG.ACC already left
      'Maria would already want to have already left him'   (Cinque 2006: 17)

To account for this difference, Cinque proposes that *volere*, unlike other restructuring verbs, can take a silent lexical OBTAIN (see also Ross 1976; Kayne 1993; Harley 2004; Harves 2008; Harves & Kayne 2012; Grano 2015). In the current context, we assume that it is not *volere* but OBTAIN which, being lexical, is associated

---

12. An anonymous reviewer points out that data like (26) are extremely interesting for the cases of defective vs. linear intervention and should be regarded and compared to other adverbs. However, the only other adverb which can be tested is 'always' (*si vorrebbe sempre aver sempre esperienze come queste* 'one would always want to always have experiences like these', Cinque 2006: 17), while all the others seem to be excluded for semantic reasons.

with its own clitic position, on a pair with lexical verbs in prototypical restructuring contexts (e.g. *voglio mangiarlo* / *lo voglio mangiare* 'I want to eat it'). It remains unexplained, however, why the clitic causee is then optionally allowed to leave the intermediate position and climb on its own onto *fare*, leaving the object stranded in its intermediate position (cf. 14).

7.   Conclusions

In this squib, we have shown that, under the correct syntactic configuration, *fare* can embed the light verb *volere*. This amounts to cases in which S_2 is not realised as a full DP (cf. i) and can be attributed to defective intervention and a ban on full DPs in (some) ECM contexts in Italian. We have also shown that when both DO and S_2 are cliticised, want-FIs exceptionally allow an accusative S_2 when the object is stranded (cf. iv), giving rise to a biclausal ECM construction. Although Italian usually disallows ECM complements to *fare*, this is also used as a repair strategy for argument licensing in PCC-contexts. Finally, we have speculated that the possibility of attaching the clitic to *volere* (cf. ii), optionally followed by climbing of one clitic only (cf. iii), arises from the additional clitic position created by the silent lexical OBTAIN selected by 'want' (Cinque 2006).

Acknowledgements

References

Belletti, Adriana & Rizzi, Luigi. 2012. Moving verbal chunks in the low functional field. In *Cartography of Syntactic Structures 7*, Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds) 129–137. Oxford: OUP.

Bonet, M. Eulàlia. 1991. Morphology after Syntax: Pronominal Clitics in Romance. PhD dissertation, MIT.

Brandi, Luciana & Savoia, M. Leonardo. 1990. Proprietà morfosintattiche e assegnazione del caso nel causativo arbëresh. *Rivista di Grammatica Generativa* 15: 29–121.

Bruening, Benjamin. 2014. Defects of Defective Intervention. *Linguistic Inquiry* 45 (4). 707–719.

Burzio, Luigi. 1981. Intransitive Verbs and Italian Auxiliaries. PhD dissertation, MIT.

Burzio, Luigi. 1986. *Italian Syntax*. Dordrecht: Reidel.
    https://doi.org/10.1007/978-94-009-4522-7

Cardinaletti, Anna & Shlonsky, Ur. 2004. Clitic positions and restructuring in Italian. *Linguistic Inquiry* 35(4): 519–557. https://doi.org/10.1162/0024389042350523

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Cinque, Guglielmo. 2003. The interaction of passive, causative, and 'restructuring' in Romance. In *The Syntax of Italian Dialects*, Cristina Tortora (ed.) 50–66. Oxford: OUP.

Cinque, Guglielmo. 2004. 'Restructuring' and functional structure. In *Structures and Beyond*, Adriana Belletti (ed.) 132–191. Oxford: OUP.

Cinque, Guglielmo. 2006. *Restructuring and Functional Heads*. Oxford: OUP.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8: 81–120. https://doi.org/10.1111/j.1467-9612.2005.00076.x

Folli, Raffaella & Harley, Heidi. 2007. Causation, obligation, and argument structure: On the nature of little v. *Linguistic Inquiry* 38: 197–238. https://doi.org/10.1162/ling.2007.38.2.197

Grano, Thomas. 2015. *Control and Restructuring*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198703921.001.0001

Guasti, Maria Teresa. 1993. *Causative and Perception Verbs: A Comparative Study*. Torino: Rosenberg & Sellier.

Guasti, Maria Teresa. 1996. Semantic restrictions in Romance causatives and the incorporation approach. *Linguistic Inquiry* 27: 294–313.

Harley, Heidi. 2004. Wanting, having, and getting: A note on Fodor and Lepore 1998. *Linguistic Inquiry* 35: 255–267. https://doi.org/10.1162/002438904323019066

Harves, Stephanie. 2008. Intentional transitives and silent HAVE: Distinguishing between want and need. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, Natasha Abner & Jason Bishop (eds) 211–219. Somerville, MA: Cascadilla Proceedings Project.

Harves, Stephanie & Kayne, Richard. 2012. Having *need* and needing *have*. *Linguistic Inquiry* 43: 120–132. https://doi.org/10.1162/LING_a_00076

Kayne, Richard.1975. *French Syntax*. Cambridge MA: The MIT Press.

Kayne, Richard. 1993. Towards a modular theory of auxiliary selection. *Studia Linguistica* 47(1): 3–31. https://doi.org/10.1111/j.1467-9582.1993.tb00837.x

Ledgeway, Adam. 2016. Italian, Tuscan, and Corsican. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 206–227. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677108.003.0014

Ledgeway, Adam. Forthcoming. 'The verb phrase: VP-related functional projections'. In *The Syntax of Italian*, Giuseppe Longobardi (ed.) Cambridge: CUP.

Manzini, M. Rita & Savoia, Leonardo M. 2007. *A Unification of Morphology and Syntax*. London: Routledge.

Marchis Moreno, Mihaela & Carolina Petersen. 2016. In defense of defective intervention: Locality effects in Romance. In *Romance Languages and Linguistic Theory, Ernestina Carrilho, Alexandra Fiéis, Maria Lobo & Sandra Pereira*, 171–190. Amsterdam: John Benjamins.

Marchis Moreno, Mihaela & Carolina Petersen. 2017. Is there any defective intervention in the syntax? In Aaron Kaplan, Abby Kaplan, Miranda K McCarvel & Edward J Rubin (eds), *Proceedings of WCCFL* 34, 355–363. Somerville, MA: Cascadilla Proceedings Project.

McGinnis, Martha. 1998. Locality in A-movement. PhD dissertation, MIT.

McGinnis, Martha. 2000. Phases and the syntax of applicatives. In *Proceedings of NELS 31*, Min-joo Kim & Uri Strauss (eds) 333–349. Amherst MA: GLSA.

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

Rizzi, Luigi. 1976. Ristrutturazione. *Rivista di Grammatica Generativa* 1: 1–54.

Rizzi, Luigi. 1986. On chain formation. In *Syntax and Semantics 19. The Syntax of Pronominal Clitics*, Hagit Borer (ed.) 65–96. New York NY: Academic Press.

Roberts, Ian. 2016. Clitics. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 786–801. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199677108.003.0048

Ross, John R. 1976. To 'have' and to not have 'have'. In *Linguistic and Literary Studies*, in Honor of Archibald A. Hill, Mohammad Ali Jazayery, Edgar C. Polomé & Werner Winter (eds) 263–270. The Hague: Mouton.

Sheehan, Michelle. 2016. Complex predicates. In *The Oxford Guide to the Romance Languages*, Adam, Ledgeway & Martin Maiden (eds) 981–993. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199677108.003.0061

Sheehan, Michelle. 2017. The Romance Person Case Constraint is not about clitic clusters. Datives, ed. by Mateu, Jaume and Pineda, Anna. Language Science Press.

Sheehan, Michelle. In progress. A parameter hierarchy for the Romance *faire-infinitif*. Ms, *Anglia Ruskin University*.

Vincent, Nigel. 1988. Italian. In *The Romance Languages*, Martin Harris & Nigel Vincent (eds) 279–313. London: Routledge.

Optional vs obligatory movement in Albanian (pseudo)-raising constructions

Giuseppina Turano

Ca' Foscari University of Venice

In this paper, I will analyse Albanian structures containing the verb *duket* 'seem'. I will show that the raising analysis given for English cannot be extended to Albanian since, in English, raising is a last resort strategy that moves a subject from an infinitive clause to a matrix finite one, in order to be Case-marked. In Albanian, the verb *duket* selects a finite clause as its complement, so, NP-movement from the embedded clause to the matrix one is not obligatory, given that the NP originates in a position where Case is assigned/checked. Thus, in Albanian, the optional movement of the NP subject will be analysed both as Topicalization and sideward movement.

Keywords: Albanian, raising, movement, case, topicalization

1.   Introduction

Within the classical Government and Binding framework (Chomsky 1981), the rule of raising from subject to subject position, at least for English, receives an account in terms of Case theory: raising verbs, such as *seem*, that select infinitival complements, allow the raising of the lower subject to the matrix clause since the external argument generated in the embedded VP needs Case but cannot receive it from a non-finite verb. This explains the contrast between (1a) and (1b):

(1)   a. *It seems John to have won a prize
      b.   John seems to have won a prize

The NP *John,* which is the subject of the verb *win,* originates from a position internal to the embedded VP but it receives no Case in (1a) since no Case is available in an infinitive construction. The ungrammaticality of (1a) is related to a violation of the Case Filter, the principle stating that an overt NP must have Case. Thus, the subject moves out of the embedded clause and raises to the matrix one, where it

gets Case from the matrix verb. Therefore, it is the Case requirement that forces the movement of the NP producing (2a) from the underlying structure (2b):

(2) a. John seems to have won a prize
    b. [$_{IP}$ John$_i$ seems [$_{IP}$ [$_{VP}$ t$_i$ to have won a prize]]]

It is forbidden to move a NP which has already received Case, as illustrated by (3):

(3) *John$_i$ seems [$_{CP}$ that [$_{VP}$ t$_i$ has won a prize]]

(3) shows that, when the lower clause exhibits tense/agreement features, raising is ungrammatical since the NP *John* is the Case-marked subject of the embedded finite verb so the Case-marking conditions are satisfied. Therefore, in English, Raising is a sort of last resort strategy when the subject cannot receive nominative Case.

In the Minimalist Program (Chomsky 2001), the Case theory has been revisited in terms of Case-agreement system. It is assumed that all NPs bear an uninterpretable Case feature that must be checked and deleted in the course of the derivation by the operation Agree. The formulation of the Case system in Chomsky (2001) involves a relation Agree which matches a probe P and a goal G having the same features: A probe P, with uninterpretable φ-features, seeks a goal G which deletes the uninterpretable φ-features of the probe, whereas the probe P checks and deletes the uninterpretable Case feature of the goal G. On minimalist assumption, the Case value assigned to a NP depends on the probe: probe T assigns and checks nominative Case, whereas *v* assigns and checks accusative Case. NP is an active goal when it has structural Case. Once features are checked and deleted, NP can no longer enter into agreement relations and is frozen in place. Under a minimalist approach, the English sentence in (1a) is ungrammatical because the nominative Case of the embedded subject has not been Case-checked in the embedded clauses since non-finite T cannot check and delete Case. The subject *John* is still active as a goal of Agree for the matrix finite T, so it raises to the matrix subject position, producing (1b).

Besides Case, the raising operation which moves the subject in (1b) is triggered by the EPP-feature of matrix T which requires the specifier position of T to be filled (Chomsky 1995, 2008). The EPP-feature of T attracts the DP with which T agrees. Therefore, what moves to the specifier of matrix T is the embedded subject.

The example in (3) is ungrammatical because the checking of the uninterpretable Case feature of the NP *John* has been satisfied by the embedded finite T and this checking operation renders the subject inactive for further agreement operations. Raising to the matrix clause is illicit. Thus, it is the Case-(checking) theory that plays a crucial role in deriving raising constructions.

In this paper, I will show that the raising analysis given for English cannot be fully extended to Albanian since the Albanian verb *duket,* that corresponds to

the English verb 'seem', gives rise to two different paradigms: optional movement, which involves an Ā position, and obligatory raising, which involves an A position.

2. Optional movement

The Albanian verb *duket* 'seem' selects a finite clause as its complement. The embedded clause is introduced by the complementizer *se/që* 'that' and appears in the indicative mood.

(4) *duket se/që Xhoni   ka fituar një çmim*
     seems that John.NOM has won   a   prize.ACC
     'It seems that John has won a prize'

The subject of the embedded clause, that displays nominative Case, agrees with the embedded indicative verb, specified for tense and agreement. So, in (4), the singular subject agrees with a singular verbal form, whereas in (5), a plural subject agrees with a plural verb form:

(5) a. *duket se/që djemtë   kanë fituar një çmim*
          seems that   boys + the.NOM have.3PL won a   prize.ACC
          'It seems that the boys have won a prize'
     b. *duket se/që ne   kemi   fituar një çmim*
          seems that   we.NOM have.1PL won   a   prize.ACC
          'It seems that we have won a prize'

In particular, the examples in (5) show that there is no agreement between the matrix verb and the embedded subject.

Interestingly, the verb *duket* cannot be followed by a clause containing a subjunctive verb, which, as is well-known, in Albanian, has replaced the infinitive mood. The use of the subjunctive to replace the infinitive in complement clauses is a phenomenon shared by Albanian, Bulgarian, Greek and Romanian and it is even observed in those languages such as Romanian and Albanian that still maintain morphological forms of infinitive. In all these languages, the subjunctive mood is used in those contexts in which Romance languages or English use the infinite. This is exemplified in (6):

(6) *Dua   të takoj   Xhonin*[1]
     Want.1SG TË meet.SUBJ.1SG John.ACC
     'I want to meet John'

---

1. Subjunctive verbs are introduced by the invariable particle *të*.

However, as the examples in (7) show, the subjunctive clause is not acceptable as the complement of *duket*:

(7) a. **duket që Xhoni   të ketë   fituar një çmim*
    seems that John.NOM TË have.SUBJ.3SG won a prize.ACC
    'It seems that John has won a prize'

   b. **duket që djemtë   të kenë   fituar një çmim*
    seems that boys + the.NOM TË HAVE.SUBJ.3PL won a PRIZE.ACC
    'It seems that the boys have won a prize'

Nor can *duket* select a clause containing an infinitive verb, which in Albanian has a periphrastic form build up by means of the preposition *për* 'for' followed by the modal particle *të*, plus an invariable participle (*për të qenë* 'to be').

(8) **duket se/që Xhoni   për të fituar një çmim*
    seems that John.NOM PËR TË won a prize.ACC
    'It seems that John has won a prize'

The impossibility to use the analytical infinitive can depend on the fact that its distribution is only restricted to purpose clauses, so it cannot be combined with *duket*.

What disallows embedding of a subjunctive clause under *duket* is somewhat mysterious.

Thus, in Albanian, the verb *duket* never takes infinitival complements nor subjunctive clauses. It only selects indicative clauses.

Despite the fact that Albanian verb *duket* selects indicative clauses, and therefore the embedded subject does not need to move into the matrix clause, it can move to the left of the verb *duket*, even if it does not agree with the matrix verb but with the embedded one.

(9) a. *Xhoni   duket se/që ka   fituar një çmim*
    John.NOM seems that has.3SG won a prize.ACC
    'John seems to have won a prize'

   b. *djemtë   duket se/që kanë   fituar një çmim*
    boys + the.NOM seems that have.3PL won a prize.ACC
    'The boys seem to have won a prize'

In (9), the subjects *Xhoni/djemtë* raise out of a tensed clause, crossing an overt complementizer, that cannot be deleted since its deletion causes ungrammaticality:

(10) a. **Xhoni   duket ka   fituar një çmim*
    John.NOM seems has.3SG won a prize.ACC
    'John seems to have won a prize'

b. **djemtë*   *duket kanë*   *fituar një çmim*
   boys + the.NOM seems have.3PL won   a   prize.ACC
   'The boys seem to have won a prize'

The NP which raises to the matrix clause originates in the embedded clause as is clear in the example in (9b) where the agreement on the verb is controlled by the plural subject *djemtë*.

To summarize, the Albanian verb *duket* shows these characteristics: a) It never takes infinitival/subjunctive complements but only selects indicative clauses; b) It does not cause obligatory NP-movement since the position where the NP originates is a position to which Case is assigned by the embedded verb; c) It allows NP-movement out of a finite clause.

In the standard GB theory, an example like (4) can be analysed as involving movement of the subject *Xhoni* 'John' to the specifier position of the embedded Infl node (here filled by the auxiliary *kam* 'have'), where nominative Case is assigned, whereas the matrix subject position remains phonetically empty. This is possible since Albanian is a pro-drop language, therefore it allows for null subjects *pro*.

Movement of the NP to the matrix SpecIP position (cf. (9)) would be problematic as the lower CP is finite and therefore the lower SpecIP position is a nominative Case-assigning position, so the subject has no reason to move into the matrix clause. Nor it is relevant, for Albanian, to satisfy the EPP of the matrix verb by raising the embedded subject since Albanian, as a pro-drop language, does not require that the specifier of the matrix Infl be lexically filled. So, it seems that in Albanian, the movement in (9) cannot be analysed as a subject to subject movement nor can be explained in terms of Case Theory. Indeed, as we will see, the raised subject occupies an Ā position distinct from SpecIP.

In a minimalist perspective, the structure in (4) represents the simplest case since the embedded C is ϕ-complete, so it selects a ϕ-complete T. The finite T *fitoj* 'win' values and deletes the nominative Case of the subject *Xhoni* 'John', moved from Spec$v$P to SpecTP, whereas the ϕ-features of the NP value and delete the ϕ-features of the probe T. Thus, the probe T agrees with the goal NP in its domain. Once the NP has moved to SpecTP, it is inactivated, with all features valued so it no longer enters into agreement relations and is frozen in place.

What about the verb *duket*? Does *duket* have ϕ-features or not? If it does, how are these checked and deleted? How is its EPP-feature satisfied? In English, the EPP-feature of the matrix T is satisfied by merging an expletive. Unlike English, Albanian has no expletives, so the matrix subject position SpecTP may remain empty. If the EPP-feature is universally strong (Chomsky 1995), one possibility is that in pro-drop languages it is satisfied via V-raising to T, according to a proposal by Alexiadou & Anagnostopoulou (1998). This allows the matrix SpecTP to remain empty.

Consider now the derivation of sentences like (9), where the embedded subject appears on the left of the matrix verb. Under Minimalism, raising of the subject out of a Case-marked position to a higher subject position should be impossible as the lower clause, in (9), is finite, so T can assign and check the nominative Case of the NP subject. Once agreement and Case assignment have been established in the lower finite clause, the subject should be inactivated and frozen in place so it should not raise further. If the subject is not an active goal for further Case checking operations, the reason for which it moves to the matrix clause must be different.

If we take as correct the claim that movement is only triggered when there is a need to check an uninterpretable feature F of an attracting head, it is necessary, for Albanian examples in (9), to determine the nature of F and which head is responsible for the NP attraction. In English, the raising operation is triggered by the EPP-feature of matrix T, which requires T's specifier position to be filled, so T attracts the NP with which it agrees. In Albanian, matrix T does not agree with the NP which appears to its left. *Duket* rather behaves like a defective probe. As such, it has not an EPP-feature and it is unable to assign nominative Case. This means that the landing site of the NP movement is not the matrix Spec TP. For Albanian, the movement of the embedded subject to the matrix clause needs to be motivated otherwise. I will assume that the position the subject reaches is actually an Ā position: a SpecTopic position. In (9), a Topic head in the matrix CP domain acts as a probe and attracts the NP subject in its specifier. Even if postulation of such features is merely stipulative, it may explain why the NP continues to be active after its Case has been checked in the embedded clause. A Topic feature makes the NP an active goal for Agree. Therefore, what we see in (9) represents a topic construction, different from the English subject-to-subject raising.

That the dislocated subject occupies a Ā position is indicated by its position with respect to other topic or focus elements occupying the left periphery of the clause.[2] In (11a), it precedes the object *çmimin*, raised in a Topic position; whereas in (11b), it precedes the object moved to a Focus position.

(11)   a.   *djemtë,   çmimin,   duket se/që e   kanë   fituar*
             boys + the.NOM prize + the.ACC seems that it.CL have.3PL won
             'As for the prize, the boys seem to have won it'
       b.   *djemtë,   ÇMIMIN   duket se/që kanë   fituar*
             boys + the.NOM prize + the.ACC seems that have.3PL won
             'THE PRIZE, the boys seem to have won'

---

2. Here, I will assume the Force-Finiteness system proposed by Rizzi (1997) to represent the left periphery of the clause.

These constructions clearly show that the subject preceding the verb *duket* is not in the matrix SpecTP position but it is in a Topic position, inside the C domain.

Even if we reverse the order of the two constituents moving the object in a position preceding that of the subject so that the subject immediately precedes the verb *duket*, the subject bears the pragmatic function like being in Topic.

(12)   a.   *çmimin,   djemtë,   duket se/që e   kanë   fituar*
           prize + the.ACC boys + the.NOM seems that   it.CL have.3PL won
           'As for the prize, the boys seem to have won it'

   b.   *ÇMIMIN   djemtë,   duket se/që kanë   fituar*
           prize + the.ACC boys + the.NOM seems that   have.3PL won
           'THE PRIZE, the boys seem to have won'

Unlike English, where raising is motivated by the lack of Case-assigning/checking in the infinitive embedded clause, in the Albanian examples in (9) it is not the Case feature but a topic feature that activates the movement of the embedded subject to the matrix clause. In Albanian, raising is Topicalization. This conclusion is consistent with recent proposals assuming that preverbal subjects of null subject languages are lexicalized in Topic. For Albanian, this conclusion has been reached by Manzini & Savoia (2007), Turano (2017) on the basis of a number of structures, such as the subjunctives and the infinitives, showing that the NP subject of these constructions never surfaces in the canonical SpecTP position but, rather, inside the C domain.

Thus, the sentences in (9) involve, at first, movement of the subject from the basic position Spec*v*P to the embedded SpecTP position, where Agree values the ϕ-features of T and checks the nominative Case of the subject. Subsequently, the NP raises to the matrix clause to check the Topic feature on the matrix C domain.

3.   Obligatory movement

In Albanian, the verb *duket* 'seem' can also trigger agreement between the matrix verb and the embedded subject:

(13)   *djemtë   duken   se/që kanë   fituar një çmim*
           boys + the.NOM seem.3PL that   have.3PL won   a   prize.ACC
           'The boys seem to have won a prize'

In (13), both the higher and the lower verb agree with the moved subject *djemtë*. The raising operation induces agreement between the matrix verb *duken* and the subject *djemtë* which also agrees with the embedded verb. (13) seems to correspond to a multiple instance of agreement, since there are two probes that need

to have their features checked and deleted and only one goal, a single NP which is involved in two checking operations. As an anonymous reviewer points out, derivations with a single goal checking the ϕ-features of multiple probes are predicted to be possible. For example, in French compound tense constructions like *Marie est arrivée* 'Mary is arrived', *Marie* checks both the participle and the finite auxiliary. This is possible when the first agreement relation between a probe (the participle) and a goal (the NP) does not check the Case feature of the goal and only the second agreement relation does (agreement with the auxiliary). Thus, a goal can remain active if its Case is not checked and check multiple probes.

The Albanian example in (13) does not correspond to that of French compound tense since, in French, the participle does not delete the nominative Case of the NP, so the NP remains an active goal, able to move to the specifier position of the auxiliary.

As for Albanian, we can imagine that the goal checks the ϕ-features of both the embedded and the matrix verbs, but when the subject moves to the embedded SpecTP position (in order to check the features of the verb *fitoj* 'win'), all the features of the NP are checked at this point, so we should not have further agreement operations between the subject and the matrix T. In particular, checking of Case by the embedded verb should render the potential goal inactive. Instead, in Albanian, the subject obligatorily moves to the left of the matrix verb, as it is shown by the contrast between (14) and (13).

(14) **duken   se/që djemtë   kanë   fituar një çmim*
      seem.3PL that boys + the.NOM have.3PL won   a   prize.ACC
      'The boys seem to have won a prize'

What is the feature which raises the subject to the matrix clause? Why is the NP not frozen in the embedded clause?

Before answering this question, let us consider what exactly the position of the subject is. It appears on the left of the matrix verb, in a position that may correspond to SpecTP or to SpecTopP. A good diagnosis to test its position is to insert a topic or a focus in the sentence. The relevant examples are provided in (15).

(15) a. **djemtë,   çmimin,   duken   se   e   kanë   fituar*
          boys + the.NOM prize + the.ACC seem.3PL that it.CL have.3PL won
          'As for the prize, the boys seem to have won it'
    b. **çmimin,   djemtë   duken   se   e   kanë   fituar*
          prize + the.ACC boys + the.NOM seem.3PL that it.CL have.3PL won
          'As for the prize, the boys seem to have won it'
    c. **djemtë,   ÇMIMIN,   duken   se   kanë   fituar*
          boys + the.NOM prize + the.ACC seem.3PL that have.3PL won
          'THE PRIZE, the boys seem to have won'

The ungrammatical examples in (15) contrast with those in (11), showing that, when there is not agreement between the subject and the matrix verb, the subject can be topicalised.

Then, the examples in (15) would seem to exclude that the subject is in a Topic position.

In addition, consider that, when the agreement between the subject and the verb is manifested on both the matrix and the embedded verb, the subject cannot even stay inside the embedded clause, in the postverbal position (16):

(16) * *duken   se/që kanë   fituar një çmim   djemtë*
          seem.3PL that have.3PL won a   prize.ACC boys + the.NOM
          'The boys seem to have won a prize'

(14) contrasts with the sentences in (4) and (5) which represent the unmarked case whereas (16) contrasts with (17), showing that when the verb *duket* does not agree with the embedded subject, the latter may remain inside the embedded *v*P/VP domain:

(17)   *duket   se/që kanë   fituar një çmim   djemtë*
          seem.3SG that have.3PL won a   prize.ACC boys + the.NOM
          'The boys seem to have won a prize'

At this point, it is clear that there is no a perfect match between structures containing *duket* (no agreement between the verb and the subject, and optional movement) and structures containing *duken* (agreement between the subject and both the matrix and the embedded verb, and obligatory movement). In structures containing *duket*, the subject can appear in different positions: in the embedded Spec*v*P, in the embedded SpecTP, in the matrix clause (SpecTopP). In structures containing *duken*, the subject cannot stay in the embedded complement: It needs to raise to the matrix clause, in a position that is strictly adjacent to that of the matrix verb.

Coming back to (13), it apparently seems to be an instance of obligatory control, since the subject of the embedded clause is interpreted as co-referential with the subject of the matrix verb. But this poses a problem since, in Albanian, like in other Balkan languages, obligatory control configurations are realized with the subjunctive mood while excluding the indicative one. See the contrast between the grammatical sentences in (18) and the ungrammatical ones in (19).

(18)   a. *Xhoni përpiqet të lexojë*
           John   tries.3SG TË read.3SG
           'John tries to read'
      b. *I Maria prospathise na diavasi*   (Terzi 1992)
           Mary   tried.3SG   NA read.3SGs
           'Mary tried to read'

      c. *Maria încearcă să scrie*   (Terzi 1992)
        Mary tries   SĂ writes
        'Mary tries to write'

(19) a. **Xhoni përpiqet se lexon*
        John tries.3SG THAT read.3SG
        'John tries to read'

    b. **I Maria prospathise oti diavase*
        Mary tried.3SG   THAT read.3SGS
        'Mary tried to read'

    c. **Maria încearcă ca scris*
        Mary tries   THAT writes
        'Mary tries to write'

In addition to requiring the subjunctive mood, obligatory control configurations do not allow the lexical complementizer, as is illustrated in the following examples, that contrast with the paradigm in (18) without a lexical complementizer:

(20) a. **Xhoni përpiqet që të lexojë*
        John tries.3SG THAT TË read.3SG
        'John tries to read'

    b. **I Maria prospathise oti na diavasi*   (Terzi 1992)
        Mary tried   THAT NA reads
        'Mary tried to read'

    c. **Maria încearcă ca să scrie*   (Terzi 1992)
        Mary tries   THAT SĂ writes
        'Mary tries to write'

The intervention of the complementizer between the two clauses causes ungrammaticality only in case of control sentences. In Albanian and Romanian, the complementizer is, in fact, compatible with the subjunctive clause but it can only be realized when the embedded subject is different from the matrix one:[3]

(21) a. *Xhoni do që të lexoni*
        John wants.3SG THAT TË read.2PL
        'John wants you to read'

    b. *As vrea ca să ramii la noi*   (Motapanyane 1991)
        would.1SG like THAT SĂ remain.2SG with us
        'I would like you to stay with us'

---

3. Greek is different since in this language, subjunctive clauses cannot be headed by a lexical complementizer (Terzi 1992).

That the complementizer plays a crucial role in control configurations is supported by another set of data coming from the Arbëresh dialect spoken at S. Nicola dell'Alto (Southern Italy), involving verbs that can select both indicative and subjunctive complements. When the control verb selects a subjunctive complement (22a)), the embedded subject is interpreted as co-referential with the matrix subject. When the control verb selects an indicative CP introduced by *se* 'that', the embedded subject is interpreted as a pronoun with free reference (22b)).

(22) a. *Beni*$_i$ *penxarin*   *pro*$_{i/*j}$ *të niset*   *menat*
     Ben.NOM thinks.3SG TË   leave.SUBJ.3SG tomorrow
     'Ben thinks to leave tomorrow'

    b. *Beni*$_i$ *penxarin*   *se pro*$_{i/j}$ *niset*   *menat*
     Ben.NOM thinks.3SG that   leave.3SG tomorrow
     'Ben thinks he/she is leaving tomorrow'

Therefore, in Albanian, sentences involving obligatory subject control predicates can only be followed by a subjunctive verbs which is not headed by the lexical complementizer, i.e. the coreference of an embedded subject with the matrix one is possible only in the absence of a complementizer.

The structure in (13) does not meet none of the two conditions: the sentential complement is an indicative clause headed by a lexical complementizer. It appears, therefore, that the sentence in (13) is significantly different from other contexts of obligatory control.

Before proceeding further, it would be interesting to examine Romanian and Greek raising constructions.

Romanian has raising both with subjunctive complements and infinitive *a*-complements, both excluded in Albanian. Starting from a sentence like (23a)) it is possible to derive both (23b)) and (23c)).[4]

(23) a. *Se pare ca studentii apreciaza acest curs*
     'It seems that the students appreciate this course'

    b. *Studentii par a aprecia acest curs*
     'The students seem to appreciate this course'

    c. *Studentii par sa aprecieze acest curs*
     'The students seem to appreciate this course'

Raising is impossible across a lexical complementizer:

(24) ** Studentii par ca sa aprecieze acest curs*
     'The students seem to appreciate this course'

---

4. The data are taken from Motapanyane (1991).

Only the example in (23a)), without NP-movement matches Albanian structures containing the verb *duket*. Thus, Albanian and Romanian are very different with respect to the data I'm analysing. The data in (23b)–(23c) contrast with Albanian that does not allow the verb *duket* 'seem' to be followed by an infinitive or a subjunctive clause. Besides that, the Romanian example in (24) contrasts with the corresponding Albanian example which is grammatical, despite the presence of the lexical complementizer.

As for Romanian, Motapanyane (1991) assumes that, in (23a), the embedded subject moves from its basic position, inside the VP, to the preverbal subject position of the embedded clause, whereas the matrix subject position is filled with a non-lexical expletive pro. In (23b), the embedded verb is infinitive, so the subject moves to the subject position of the matrix clause. In (23c), the NP raises to the matrix subject position, but it is coindexed with an expletive pro in the embedded subject position and with the trace of the NP, in the base-generated SpecVP position. According to Motapanyane, (25) is the appropriate representation of NP-raising from finite clause.

(25) NP_i……. $E\text{-}pro_i$………t_i

The derivation in (25) observes the Binding Conditions on pronouns since, according to the author, in (25), the subjunctive particle prevents government from matrix verb on the structural subject position in the embedded clause. E-pro is in a different governing category.

But Motapanyane's idea is unsatisfactory since the derivation in (25) does not explain why the NP cannot cross the complementizer (cf. (24)). For English sentences like (3), the ungrammaticality is related to an ECP violation (Chomsky 1986): the trace of the moved NP is not properly governed.

Being E-pro a pronominal element, it should not be subjected to ECP, so, the intervention of the complementizer should not interfere causing ungrammaticality, contrary to fact. In order to solve this problem, Motapanyane (1991) assumes that what is relevant in (24) is again Binding Theory: the subjunctive CP cannot ensure an opaque domain for the embedded clause. The complementizer is unable to prevent government from the matrix verb. Therefore, (24) is ungrammatical because the Principle B of the Binding Theory is violated. The problem is circular: in (23c), the subjunctive particle is able to separate the government domains in matrix and embedded clause, while the same element in (24), plus the lexical complementizer, cannot ensure an opaque domain for the embedded clause. Motapanyane, aware of the problem, solves it problem resorting to the ECP: in (24), which has the derivation in (26), ECP is violated at the level of chain-links.

(26) NP_i…….ca….$E\text{-}pro_i$………t_i

Motapanyane assumes that, in (26), both E-pro and t must be head governed. The matrix verb can govern the embedded subject position across an IP subjunctive but not across a CP. The lexical complementizer prevents the verb to govern the E-pro. The ungrammaticality of (24) is formulated in terms of ECP.

Now consider Greek (quoted from Motapanyane 1991). The behaviour of the verb *fenete* 'seem' is exemplified by the following sentences:

(27) a.   *Fenete oti meriki   den diarazoun*
           seems that some.3PL not study
           'It seems that some people do not study'

  b.   *Meriki   fenete oti   den diarazoun*
         some.3PL seems that not study
         'Some people seem not to study'

  c. * *Meriki   fenonde oti   den diarazoun*
         some.3PL seem.3PL that not study
         'Some people seem not to study'

  d.   *Meriki   fenonde na ehoun agorasi afta   ta   vivlia*
         some.3PL seem.3PL NA have   bought these the books
         'Some people seem to have bought these books'

Greek allows NP-movement only in absence of agreement (cf. (27b) vs (27c)); it allows NP-movement across a complementizer (27b); it allows NP movement from a subjunctive clause and in this case it presents agreement features on the matrix verb (27d). Only (27a) and (27b) match Albanian structures with *duket*. Unlike Albanian, the Greek structure (27c) where both the matrix and the embedded verbs are inflected is ungrammatical. Motapanyane assumes that in (27b), the NP is left-dislocated (presumably in SpecCP), whereas (27c) is ungrammatical since the presence of a complementizer does not allow for the formation of a correct Th-chain. (27d) is analysed in a pair with the corresponding Romanian sentence in (23c), which has the representation in (25).

Thus, the comparison between Albanian, Romanian and Greek has shown that only in Albanian it is possible to extract a NP from an embedded CP-indicative complement.

Now, let us come back again to (13), that exhibits the characteristics of a structure with obligatory control since the single subject *djemtë* refers at the same time to both the matrix and the embedded verb, irrespective of the presence of the complementizer. The immediate questions are: how would be control effects realized through a CP clause boundary, considering that in Albanian the lexical complementizer has a blocking effect (cf. (19a))? Is it possible to analyse the structure in (13) as an obligatory control sentence?

Since in Albanian, control effects are attested in structures containing inflected subjunctive verbs, Turano (1993, 1995) assumes that control effects are manifested in presence of *pro,* while excludes the possibility that obligatory control sentences involve a PRO subject.

In particular, following Manzini (1983) and Borer (1989), Turano assumes that in Albanian, control reduces to binding. The author assumes that in obligatory control sentences, the T node is anaphoric, so that it needs to be bound by a referential Tense or by an argument of the matrix clause. In order to fix its reference, T must move to the C position. When T does not raise to C because C is already occupied by the complementizer *se*, binding of the anaphoric T from the matrix clause is blocked.[5] Thus, it is a rather undesirable conclusion to assume that, in the case of (13), the coreference of the embedded subject with the matrix one is possible also in presence of a lexical complementizer. The complementizer should prevent movement of the embedded verb to C and, as a consequence, the embedded T would not be bound by the matrix verb.

In light of these considerations, I will propose, although tentatively, an alternative analysis. I assume that sentences like (13) can be derived under a sideward movement analysis as proposed by Boeckx (2001) for donkey-sentences. Following the analysis of parasitic gaps developed by Nunes (1995, 2000), Boeckx takes the two occurrences of *a donkey* in (28) to be two non-distinct copies of the same element, related by sideward movement:

(28)   if John owns a donkey^i, he beats a donkey^i

According to Boeckx, the initial numeration contains two copies of the NP. Upon building the matrix clause, *a donkey* is copied and merged, in a parallel derivation, into the protasis: the computational system copies a given element $\mathbb{V}$ of a syntactic object K and merges $\mathbb{V}$ with a syntactic object L, which has been independently assembled and is unconnected to K (Boeckx 2001: 14).

(29)   a. $[_K \dots \mathbb{V}^i \dots]$ $\mathbb{V}^i$ <> $[_L \dots]$
                                  $\lfloor\_$ copy $\_\rfloor$

          b. $[_K \dots \mathbb{V}^i \dots]$ $[_M \mathbb{V}^i [_L \dots]]$

Following this analysis, I will propose that, in (13), a first copy of the NP subject is merged in the embedded *v*P (for φ-theory) whereas a second one is merged in SpectTP (for matching of φ-features). In the final step, a third copy is created

---

5. I assume that control effects are imposed by the lexical properties of control verbs. Therefore, the subjunctive T is coreferential with a matrix verb only when it has certain selected properties.

in the matrix clause when the embedded TP merges with CP. This copy is pronounced here, whereas the other copies are deleted. Thus, the first checking operation between the probe T and the goal NP inside the embedded TP produces erasure of the uninterpretable features of the embedded verb but nominative Case is not deleted. Therefore, the checking operation continues in the matrix clause when the higher copy of the subject matches with the features of the matrix verb. At this point the nominative Case of the NP is checked and deleted. Therefore, (13) is the result of a sideward movement of the NP subject to the matrix SpecTP position with deletion of the copies left by movement.

4.   Conclusions

In this paper, I have analysed Albanian structures with the raising verb *duket*, that corresponds to the English verb *seem*. I have shown that the raising analysis given for English cannot be extended to Albanian, since in English, raising is a last resort strategy that moves a subject from an infinitive clause to a matrix finite one, in order to be Case-marked. In Albanian, the verb *duket* selects a finite clause as its complement, rather than a nonfinite complement. Then, NP-movement from the embedded clause to the matrix one is not obligatory, since the position where the NP originates is a position to which Case is assigned and checked. Besides that, in Albanian, raising takes place from a finite indicative clause introduced by the complementizer *se/që* 'that', a clear indication that the embedded complement clause is a full CP projection. Since Case is not responsible for movement of the subject to the matrix clause, in Albanian, the optional movement of the NP is analysed as Topicalization whereas the obligatory movement of the NP is analysed as sideward movement.

Acknowledgements

References

Alexiadou, Artemis & Anagnostopoulou, Elena. 1998. Parametrizing Agr: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory* 16(3): 491–539.
https://doi.org/10.1023/A:1006090432389

Boeckx, Cedric. 2001. Gaps and donkeys: Preliminary remarks. In *Proceedings of the first Harvard-MIT Student Conference in Language Research*. Cambridge, MA: The MIT Press.

Borer, Hagit. 1989. Anaphoric Agr. In *The Null Subject Parameter*, Osvaldo Jaeggli & Ken Safir (eds), 69–110. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-009-2540-3_3

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

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

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

Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*. Cambridge MA: The MIT Press.
https://doi.org/10.7551/mitpress/9780262062787.003.0007

Dobrovie-Sorin, Carmen. 1991. The Syntax of Romanian. Comparative Studies in Romance. Ms, Université de Paris 7.

Manzini, M. Rita. 1983. On Control and Control Theory. *Linguistic Inquiry* 14: 421–446.

Manzini, M. Rita & Savoia, Leonardo. 2007. *A Unification of Morphology and Syntax. Investigations Into Romance and Albanian Dialects*. London: Routledge.

Motapanyane, Virginia. 1991. Theoretical Implications of Complementation in Romanian. PhD dissertation, University of Geneva.

Nunes, Jairo. 1995. The Copy Theory of Movement and Linearization of Chains in the Minimalist Program. PhD dissertation, University of Maryland.

Nunes, Jairo. 2000. Sideward movement and the linearization of chains. Ms, Unicamp.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, Liliane Haegeman (ed.), 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7

Terzi, Arhonto. 1992. PRO in Finite Clauses. A Study of the Inflectional Heads of the Balkan Languages. PhD dissertation, University of New York.

Turano, Giuseppina. 1993. Subjunctive constructions in Arbëresh and standard Albanian. *Rivista di Grammatica Generativa* 18: 101–133.

Turano, Giuseppina. 1995. *Dipendenze sintattiche in albanese*. Padova: Unipress.

Turano, Giuseppina. 2017. Modal particles in Albanian Subjunctive, infinitive and Supine constructions: Presence vs absence of clitic climbing. *Quaderni di Linguistica e Studi Orientali* 3: 361–376.

Clitics and pronouns
from a theoretical perspective

Clitic stress allomorphy in Sardinian

Laura Bafile and Rosangela Lai

University of Ferrara / University of Florence

Sardinian displays stress shifts under cliticisation with imperative and gerund verb forms. Stress shift is related to the type and number of clitics associated to the host. Across the range of dialectal variation, three different stress shift patterns are attested. We will argue that Sardinian data supports the approach whereby stress shift variation cannot be regarded either as the result of purely prosodic rules or as the consequence of different syntactic feature-checking properties of the clause. The analysis here proposed accounts for stress placement as an allomorphy that is partly determined by phonological conditions.

Keywords: clitics, allomorphy, stress shift, Sardinian

1.   Introduction

Romance cliticisation is a relevant issue for phonological studies inasmuch as enclitic attachment, in some Romance languages, involves the metrical reshaping of the sequence Verb+clitic(s) (henceforth 'clitic group'). This phenomenon, commonly defined 'stress shift', has been treated within different models. In purely phonological accounts, stress shift is attributed to the same kind of rules that regulate the prosodic structure of the word and that are deemed to extend to enlarged domains. According to a different hypothesis (Ordóñez & Repetti 2006, 2014), stress shift phenomena actually reflect the existence of three categories of pronouns, namely strong pronouns, weak pronouns and clitics, each one with its own syntactic and phonological properties. In some languages verbs in non-finite forms cannot attract clitics; as a result, so called 'stressed enclitics', i.e. postverbal pronouns associated to imperatives, gerund and infinitive, actually belong to the class of weak pronouns.

   In this paper, we will take into account in some detail stress patterns of clitic groups and stress shift phenomena in Sardinian dialects. This analysis may contribute to a broad picture of the phonological aspects of cliticisation in Romance

languages, as well as to the investigation of the role that different grammar modules play in regard to this phenomenon.

The paper is organised as follows: In Section 2, we focus on some elements of previous proposals that are specifically relevant to our discussion. We then analyse stress patterns of clitic groups in Sardinian dialects (Section 3) and discuss specific data that are puzzling for existing analyses (Section 4). Section 5 takes briefly into account preverbal clitics. In Section 6, we draw some conclusions and look at suggestions for future research.

2. Approaches to enclitic stress phenomena

Two main approaches, belonging to completely different frameworks, have been proposed to account for stress shift phenomena caused by cliticisation that can be observed in Romance languages like Sardinian, Catalan, Occitan, Neapolitan and Lucanian dialects. Phonological accounts generally rely on the idea that enclitic stress shift is the result of metrical principles and parameters that hold in the phonological word and that extend to enlarged domains. The alternative explanation (Ordóñez & Repetti 2006, 2014) assumes the existence of three categories of pronouns, namely strong pronouns, weak pronouns and clitics, each one with its own syntactic and phonological properties. In this view, stress shift phenomena actually reflect the lexical-syntactic content of pronouns.

2.1 Phonological analyses

Phonological accounts of stress shift in clitic groups consider this phenomenon as resulting from assignment of metrical structure (or, in notationally different terms, prosodic structure) to domains enlarged by clitic adjunction. Work on stress structure of enlarged domains, mostly, but not exclusively, consisting of clitic groups, has been carried out in the framework of Halle and Vergnaud's (1987) stress theory (cf. Halle 1990; Halle & Kenstowicz 1991; Kenstowicz 1991). In this perspective, stress shift actually consists in the reassignment of metrical structure after the lexical cycle. Postcyclic metrification preserves the pre-existent structure, in respect of a principle of Structure Preservation (Steriade 1988), and assigns a new stress, which becomes the main stress, on the postcyclic section of the enlarged domain. In case of clash, the lexical stress is deleted.

In phonological accounts, the variation between stress-shift and non-stress-shift languages directly depends on a parametric choice, variously formulated, that allows or prevents stress reassignment after word level. The various patterns observable in stress-shift languages correspond to subtler parametric options. In

metrical accounts (cf. Kenstowicz 1991), this variation results from the interaction of two variables in the algorithm, i.e. structure preservation and extrametricality. In (1) we illustrate the mechanism triggered by the adjunction of a single clitic. Given extrametricality of the final syllable, the strict application of structure preservation causes stress shift, as in *por'talə* 'bring it'. On the other hand, a looser application allows the final extrametrical syllable of a paroxytonic base to be incorporated in the preceding foot, as in *'portalə*, while the adjoined clitic becomes, in turn, extrametrical ('< >' indicate extrametrical content):

(1)   *Strict Structure Preservation*
      (por)<ta > = lə → (por)(ta)<lə > [por'talə].
      Loose Structure Preservation:
      (pɔr)<ta > = lə → (pɔr.ta)<lə > ['pɔrtalə].

In a different account, by a loose interpretation of structure preservation, stress reassignment operates as a repair strategy: a new stress is assigned to the clitic group only if the attached syllables cannot be directly incorporated into the existing metrical structure. This analysis has been applied to Neapolitan by Bafile (1993, 1994), in which stress is reassigned, i.e. shifts, only when two or more clitics are attached. The proposal is consistent with the fact that proparoxytonic words are common in this language, while preantepenultimate stress is generally not allowed. Bafile's proposal excludes extrametricality from the representation:

(2)   *Loose Structure Preservation*
      (por.ta) = lə → (por.ta.lə) ['portalə]
      (por.ta) = tə = lə → (po.rta) (tə.lə) [pɔrta'tillə]

In a similar vein, in the analyses developed in the framework of Prosodic Phonology, stress patterns depend on the way clitics are integrated into the prosodic hierarchy. For example, in Peperkamp (1997) different patterns in Italian stress-shift languages are accounted for as the direct effect of either adjunction or incorporation into the prosodic word. Based on the assumption that feet are maximally binary, Paperkamp's proposal is that by adjunction, when a single clitic is attached it is not metrified and is included in a recursive level of the prosodic word; as a consequence, no stress shift takes place (e.g. Neapolitan *'vinnələ* 'sell it'). By incorporation, a single clitic is inserted in the prosodic word, with the result that a binary foot is built on the right edge of the word (e.g. Lucan *vən'nillə* 'sell it'). If two clitics are inserted, a binary foot is built on the right edge in any case.

The weakness of exclusively phonological, algorithmic accounts of clitic stress shift was acknowledged in some early papers on this subject. For example, Bafile (1993, 1994) argues that disyllabic forms like *-'tillə* (DAT.2SG = ACC.3SG.M) *-'tennə* (DAT.2SG = LOC/PART) in Neapolitan should be analysed as allomorphs

corresponding to clitic clusters, rather than as plain sequences of clitics such as *te = lo*, *te = ne*.

Subsequent work on clitics and stress shift has highlighted several weak points in the empirical adequacy of purely phonological analyses. The most common criticism is that, as a matter of fact, in most stress-shift languages, allegedly illegal stress patterns are allowed in some contexts. One example is the case of Lucanian where, in sequences Verb+Cl, penultimate stress may alternate with antepenultimate, e.g. *narrá = la* 'tell it (f.)' vs. *nárra = mi* 'tell me' (Ordóñez & Repetti 2006) and, more in general, antepenultimate stress is allowed in words as well as in Noun+Cl sequences, e.g. *lávənə* 'they wash', *frátə = mə* 'my brother' (Manzini & Savoia 2017). Moreover, some languages show in clitic groups a pattern that is otherwise less common in the lexicon, i.e. stress on the final syllable, e.g. *vindi = rú* 'sell it' in Ligurian, *porto = ma = rú* 'bring it to me' in Gascon (cf. Pescarini forthcoming a; Ordóñez & Repetti 2006).

2.2 The 'weak-pronoun' analysis

The failure of phonological explanations is the starting point of the proposal by Ordóñez & Repetti (2006, 2014). Building on the idea of Cardinaletti & Starke (1999) that pronouns divide into three different categories, Ordóñez & Repetti assume that the pronouns involved in cliticisation may be either true clitics or weak pronouns. They argue, then, that the variation in Romance languages concerning stress shift depends on this difference. The fundamental assumption is a phonological one, stating that clitics are unstressed and that stressed elements cannot be clitic; therefore, the latter must be regarded instead as weak pronouns. Their proposal is that when the verb is in the position T of the syntactic tree, as in finite clauses, it can associate with clitics. When the verb is in C, as in imperative clauses, in some languages it cannot associate with clitics; as a consequence, pronouns are attracted by the lower position *v*, which can only associate with weak pronouns. Subtler variation concerning the form of postverbal pronouns is attributed to parametric conditions that regulate the attractive capacity of *v*. In sum, according to Ordóñez & Repetti, preverbal pronouns are true clitics while stressed postverbal pronouns must rather be regarded as weak pronouns.

Both the strong-weak-clitic tripartition and the weak-pronoun analysis have attracted criticism (see Manzini 2014; Pescarini forthcoming a, forthcoming b). In essence, it is argued that, on the one hand, diagnostics for the definition of classes of pronouns do not hold cross-linguistically and are often contradictory and that, on the other hand, alternative accounts are available. More specifically, according to Manzini & Savoia (2017), the weak-pronoun analysis of stressed enclitics is problematic for empirical reasons. They observe that between the C position,

taken by the verb in imperatives, and the *v* position, that attracts weak pronouns in Ordóñez & Repetti's analysis, no kind of material is ever inserted. Given the considerable syntactic space between the two positions, we could expect to find intervening material, e.g. adverbs, which however is not the case.

In Section 5, we will show data concerning preverbal cliticisation in the Sardinian varieties we are focusing, and argue that they raise a further problem for the weak-pronoun analysis.

3.   Enclitic stress in Sardinian

On the whole, Sardinian dialects form a consistent picture as far as lexical stress is concerned, sharing the general properties of Italo-Romance stress system. However, considerable variation is observed as far as enclitic stress is concerned.

3.1   General properties of lexical stress

Traditionally, the Sardinian language group is divided in two macro-areas: Campidanese (southern half of Sardinia) and Logudorese (central-northern Sardinia). In both areas, a remarkable internal variation affects every level of the language (phonology, morpho-syntax and lexicon). A transitional area (Barigadu, Lower Barbagia and central-northern Ogliastra), placed between Campidanese and Logudorese, shows interesting peculiarities, including stress shift patterns (Pisano 2007: 51).

Generally speaking, Sardinian lexical stress system is quite homogeneous (Wagner 1941: 25ff; Pittau 1972: 20ff; Bolognesi 1998: 65ff.). Stress may fall on one of the last three syllables, although most words have penultimate or, less frequently, antepenultimate stress. In all Sardinian varieties, oxytone words are usually repaired with the insertion of a paragogic vowel or (rarely) of a whole syllable, e.g. Italian *caffè* → Campidanese *caffèi*; Catalan *vosté* → Campidanese *vostéti* 'you (polite form)' (Wagner 1941: 25).

Pittau (1972: 21) emphasises that in Nuorese (a Logudorese variety) the near totality of words are paroxytone or proparoxytone, while pre-proparoxytone words are not attested. All the exceptions result from cliticisation. In Logudorese, when a single clitic is added to proparoxytone imperatives, e.g. *bókina* 'call (2SG. IMP.)', stress is on the fourth syllable from the right, e.g. *bókina = lu* 'call him!' (Pittau 1972: 21).

In Campidanese, on the other hand, we find a series of stressed monosyllables that, historically, resulted from the deletion of an intervocalic voiced stop: IUGU> /dʒu/ 'yoke', NIVE> /ni/ 'snow', DIE> /di/ 'day', FABA> /fa/ 'broad bean', CRUDU> /kru/

'raw' (Wagner 1941: 25). According to Bolognesi (1998: 65, 292), these monosyllables, when in phrase final position, are repaired by the epenthetic syllable [ɾi] thus resulting in a trochaic form. See (3) below (adapted from Bolognesi (1998: 292).

(3)   /ssa di/ → [sa ˈdiɾi] 'the day' (Southern Campidanese – Sestu variety)

3.2   Enclitic stress patterns

As far as clitic stress is concerned, Sardinian dialects show quite diverse patterns. Despite the fine-grained dialectal variation, three patterns of enclitic stress can be identified, corresponding to the Logudorese, the Campidanese and the dialects of the Transitional area.

In Logudorese, the attachment of a single clitic does not affect verb stress; for example, with the host *bátti* 'bring (2sg.imp)', *báttimi* 'bring to me'. On the contrary, with two or three clitics, stress shifts to the penultimate syllable of the string, e.g. *battimíla* 'bring it (f) to me'.

In Campidanese, with single clitics, stress falls either on the final or on the penultimate syllable. Strings with first and second person clitics have final stress, e.g., with *bétti* 'bring (2sg.imp)', *bettimí* 'bring (sth) to me', while, with third person clitics, stress falls on the penultimate syllable, e.g. *bettídda*. The adjunction of two or more clitics induces stress shift to the penultimate syllable, e.g. *bettimídda* 'bring it (f) to me'.

Generally speaking, within the dialects of the Transitional area, there is strong phonological and morpho-syntactic variation. However, as far as stress-shift is concerned, the Transitional area is internally consistent (Pisano 2007: 51). In these varieties, cliticisation always creates paroxytone forms, independently of the number of clitics, e.g. *battími* 'bring (sth) to me', *battimídda* 'bring it (f) to me'.

In the next sections we consider in some detail the three Sardinian enclitic stress patterns.

3.2.1   *Logudorese*
In Logudorese, stress shift takes place only when two or three clitics are attached, always producing penultimate stress (Wagner 1941: 24):

(4)   Logudorese
      a. *Single enclitic: No stress shift*
      ˈmaniɣalu 'eat it/him'   eat-imp.2sg = acc.3sg.m
      ˈnarami 'tell me'   tell-imp.2sg = dat.1sg
      ˈnaralu 'tell it (m)'   tell-imp.2sg = acc.3sg.m
      ˈnarali 'tell him'   tell-imp.2sg = dat.3sg
      ˈkɔlabbi 'come by there'   come.by-imp.2.sg = loc

'bokinalu 'call him'   call-IMP.2SG = ACC.3SG.M
is'kuzami 'excuse me'   excuse-IMP.2SG = ACC.1SG
'battimi 'bring (sth) to me'   bring-IMP.2SG = DAT.1SG
'battila, 'battilu 'bring it (f/m)'   bring-IMP.2SG = ACC.3SG.F/M
'battikke 'bring from there'   bring-IMP.2SG = LOC

> Data from Pittau (1972: 20, 82, 85); Blasco Ferrer (1984: 261); Pisano
> (2007: 51); Lai (2017: 189).

b. *Two or three clitics: Penultimate stress*

Two clitics

maniɣa'ðilu 'eat it/him.'   eat-IMP.2SG = DAT.2SG = ACC.3SG.M
pikka'tiɳɖe 'take some (for yourself)'   take-IMP.2SG = DAT.2SG = PART
bazeb'bɔŋke 'go away from there'   go- IMP.2SG = DAT.2SG = LOC
nara'βilu 'tell it to him'   tell-IMP.2SG = DAT.3SG = ACC.3SG.M
bessi'mikke 'make way (for me)'   exit-IMP.2SG = DAT.1SG = LOC
batti'kɛla, batti'kelu 'bring it (f/m)   bring-IMP.2SG = LOC = ACC.3SG.F/M
there'
batti'mikke 'bring (sth) there for me' bring-IMP.2SG = DAT.1SG = LOC

Three clitics

porrimiɳ'ɖɛla 'hand it (f) out to me   pass-
from there'   IMP.2SG = DAT.1SG = LOC = ACC.3SG.F
buffati'kelu 'drink it (to your health)' drink-
                                        IMP.2SG = DAT.2SG = LOC = ACC.3SG.M
bogademi'kɛla 'get it (f) out of my   pull-
way'   IMP.2SG = DAT.1SG = LOC = ACC.3SG.F
battimi'kɛla, battimi'kelu 'bring it   bring-
(f/m) there for me'   IMP.2SG = DAT.1SG = LOC = ACC.3SG.F/M

> Data from Pittau (1972: 85); Blasco Ferrer (1984: 257, 1986: 114);
> Lai (2017: 189)

3.2.2   *Transitional area*

Varieties of Transitional area have generalised stress shift to the penultimate syllable (Wagner 1941: 24):

(5)   Transitional area
      a. *Single enclitic: Penultimate stress*
      na'rami 'tell me'   na'rami 'tell me'
      na'rami 'tell me'   tell-IMP.2SG = DAT.1SG
      affer'raɖɖu 'grab it/him'   grab-IMP.2 G = ACC.3SG.M
      kor'kaði 'lay yourself down'   lay.down-IMP.2SG = ACC.2SG
      set'tsiði 'sit down'   sit-IMP.2SG = ACC.2SG
      pen'sakke 'think about it'   think-IMP.2SG = LOC

settʃio'rosi 'sit down' sit-IMP.2PL = ACC.2PL

isku'zami 'excuse me' excuse-IMP.2SG = ACC.1SG

fa'emmi 'do (sth) for me' do-IMP.2SG = DAT.1SG

bat'timi 'bring (sth) to me' bring-IMP.2SG = DAT.1SG

bat'tiɳɖe 'bring (sth) from there' bring-IMP.2SG = LOC

bat'tiɖɖa 'bring it (f)' bring-IMP.2SG = ACC.3SG.F

Data from Blasco Ferrer (1988: 112–113); Pisano (2007: 51); Lai (2017: 189).

b. *Two or three enclitics: Penultimate stress*

Two clitics

bae'tiɳɖe 'go away' go-IMP.2SG = ACC.2SG = LOC

dona'miɖɖu 'give it to me' give-IMP.2SG = DAT.1SG = ACC.3SG.M

batti'miɖɖa, batti'miɖɖu bring-IMP.2SG = DAT.1SG = ACC.3SG.F/M
'bring it (f/m) to me'

battiɳ'ɖɛɖɖa, battiɳ'ɖɛɖɖu bring-IMP.2SG = LOC = ACC.3SG.F/M
'bring it (f/m) from there'

batti'miɳɖe 'bring (sth) to bring-IMP.2SG = DAT.1SG = LOC
me from there'

Three clitics

porrimiɳ'ɖɛɖɖa 'hand it (f) pass-IMP.2SG = DAT.1SG = LOC = ACC.3SG.F
out to me from there'

pixaðik'keɖɖu 'take it (m) take-IMP.2SG = DAT.2SG = LOC = ACC.3SG.M
for yourself from there'

battimiɳ'ɖɛɖɖa, bring-IMP.2SG = DAT.1SG = LOC = ACC.3SG.F/M
battimiɳ'ɖɛɖɖu 'bring me it
(f/m) from there'

Data from Blasco Ferrer (1988: 112–113); Lai (2016: 141).

3.2.3 *Campidanese*

The case of the Campidanese stress shift pattern is the most complex. As shown in (6a) below, sequences with clitics of first and second person singular (i.e. *-mi* [mi], *-di* [ði]), first and second person plural, (i.e. *-sí* [zi] ~ *-nosí* [nozi], *-sí* [zi] ~ *-osí* [ozi])[1] have different stress compared to sequences with the third person singular or plural clitics (i.e. *-dda/u/i* [ɖɖa/u/i], *-ddas/us/is* [ɖɖas/us/is]). The former have final stress, while the latter have penultimate stress see (6b). With partitive and locative clitics, stress is on the penultimate syllable as well see (6b), and the same holds with all the sequences with two or three clitics (6c). Data in (6a), (6b), (6c), (6d) are from the variety of Tertenía (Northern-Eastern Campidanese).

---

1. The voiceless alveolar sibilant is realized [z] in intervocalic position.

(6)   Campidanese
      a. *Single Clitic (1st or 2nd Person): Final stress*
      Singular
      nara'mi 'tell me'   tell-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1sɢ
      aβeri'mi 'open (sth) for me'   open-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1sɢ
      piɣa'mi 'take (sth) for me'   take-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1sɢ
      korka'ði 'lay yourself down'   lay.down-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.2sɢ
      setʧi'ði 'sit down'   sit-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.2sɢ
      piɣa'ði 'take (sth) for yourself   take-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.2sɢ
      Plural
      kompra'zi ~ komprano'zi 'buy (sth) for us'   buy-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1ᴘʟ
      aβeri'zi ~ aβerino'zi 'open (sth) for us'   open-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1ᴘʟ
      betti'zi ~ bettino'zi 'bring (sth) to us'   bring-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1ᴘʟ
      kompra'zi ~ kompraio'zi 'buy (sth) for   buy-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.2ᴘʟ
      yourselves
      aβeri'zi ~ aβerio'zi 'open (sth) for yourselves'   open-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1ᴘʟ
      betti'zi ~ betteio'zi 'bring (sth) for yourselves'   bring-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.1ᴘʟ

      b. *Single Clitic (3rd Person Accusative, Dative and Adverbial clitics):*
      Penultimate stress
      Accusative
      pap'padɖa 'eat it (f)'   eat-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.3sɢ.ғ
      aβe'rɛdɖa 'open it (f)'   open-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.3sɢ.ғ
      la'madɖa 'call her'   call-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.3sɢ.ғ
      imbuʃʃadɖa 'cover it (f)'   cover-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.3sɢ.ғ
      appes'sadɖa 'fold it (f)'   fold-ɪᴍᴘ.2sɢ = ᴀᴄᴄ.3sɢ.ғ
      bet'tɛdɖa 'bring it (f)'   bring-ɪᴍᴘ.2sɢ = ʟᴏᴄ = ᴅᴀᴛ.2sɢ
      Dative
      na'redɖi 'tell him'   tell-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.3sɢ.ᴍ
      can'tedɖi 'sing for him'   sing-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.3sɢ.ᴍ
      aβe'redɖi 'open (sth) for him'   open-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.3sɢ.ᴍ
      kom'pradɖi 'buy (sth) for him'   buy-ɪᴍᴘ.2sɢ = ᴅᴀᴛ.3sɢ.ᴍ
      Partitive and locative clitics
      bet'tɛndi 'bring (sth) from there for   bring-ɪᴍᴘ.2sɢ = ʟᴏᴄ = ᴅᴀᴛ.2sɢ
      yourself'
      pi'ɣandi 'take some'   take-ɪᴍᴘ.2sɢ = ᴘᴀʀᴛ
      kom'prandi 'buy some'   buy-ɪᴍᴘ.2sɢ = ᴘᴀʀᴛ
      se'ɣandi 'carve some'   carve-ɪᴍᴘ.2sɢ = ᴘᴀʀᴛ

c. *Two or three clitics: Penultimate stress*
Two

betti'midɖa 'bring it (f) to me'	bring-IMP.2SG = DAT.1SG = ACC.3SG.F
peza'midɖa 'bring it (f) for me'	bring-IMP.2SG = DAT.1SG = ACC.3SG.F
betti'mindi 'bring (sth) to me from there / bring some to me'	bring-IMP.2SG = DAT.1SG = LOC/PART
pɔni'ðindi 'put some (somewhere) for yourself'	put-IMP.2SG = DAT.2SG = PART

Three

pɔnimin'dɛdɖa, pɔnimin'dedɖu 'it (f/m) there for me'	put-IMP.2SG = DAT.1SG = LOC = ACC.3SG.F/M
pezamin'ʧɛdɖa 'bring it (f) there for me'	bring-IMP.2SG = DAT.1SG = LOC = ACC.3SG.F
piɣamin'dedɖu 'take it (m) from there for me'	take-IMP.2SG = DAT.1SG = LOC = ACC.3SG.M

The clitic forms *-nosí* and *-osí*, respectively 1PL and 2PL clitics (Blasco Ferrer 1986: 111), are also listed in (6a). Nowadays, these forms are only ever found in the most conservative areas of the Campidanese domain, i.e. Ogliastra and Lower Barbagia (Northern Campidanese). Younger speakers usually replace them with the form *-sí*. Both *-sí* and the older forms *-nosí*, *-osí* are grammatical for Campidanese speakers. Notice that *-sí* is also the third person singular and plural reflexive enclitic, which associates with gerund forms.[2] In Campidanese, *-sí* always bears stress when it is the only clitic in the sequence; the same holds for the bisyllabic forms of first and second person plural clitics *-nosí, -osí,* that also have final stress:

(7)   *Syncretic Form* -si (*final stress*)
Imperative + 1PL/2PL

kompra'zi 'buy (sth) for us/you'	buy-IMP.2SG = DAT.1PL/2PL
aβeri'zi 'open (sth) for us/you'	open-IMP.2SG = DAT.1PL/2PL
betti'zi 'bring (sth) to us/you'	bring-IMP.2SG = DAT.1PL/2PL

Gerund + Third person reflexive

askurtɛndu'zi 'listening to oneself'	listen-PROG = REFL
pettenɛndu'zi 'combing oneself'	comb-PROG = REFL
samunɛndu'zi 'washing yourself'	wash-PROG = REFL

---

2. Gerund is another mood (besides imperative) that allows for enclitics in Sardinian. The stress patterns that we have described for imperative, and their geographic distribution, carry over the gerund case without any changes.

An indirect evidence that these sequences take the main stress on the final syllable is in the fact we mentioned above that in many varieties of Campidanese a vowel or a syllable is inserted as a repair strategy at the right edge of oxytonic words (Pisano 2007: 52). Lower Barbagia (Northern Campidanese) employs a paragogic vowel [a], e.g. settsi'ði[a] 'sit down', while in southern areas the syllable [ɾi] is inserted instead, e.g. skuza'mi[ɾi] 'excuse me' (data adapted from Pisano 2007: 52). As these examples show, the same repair strategy is applied in oxytonic words as well as in clitic strings.

4.   Discussion

At first sight, a robust phonological motivation could be found in the enclitic stress system of the Transitional area, given that in this variety stress always falls on the penultimate syllable, independently of the composition of the enclitic string. Stress reassignment in enlarged domains, it could be claimed, is consistent with the fundamental, optimal metrical constituent, i.e. the binary, left-headed foot. As far as Logudorese is concerned, a phonological motivation could also be available, inasmuch as the enclitic stress pattern is here compatible with the analysis we mentioned above about Neapolitan, whereby stress shift only applies when two or more clitics are attached, to repair ungrammatical sequences of three unstressed syllables (cf. Bafile 1993, 1994).

On the other hand, phonological accounts would result in obvious failure when applied to Campidanese, given that no phonological regularity emerges from the data of that variety. When one clitic is attached, stress falls either on the final or on the penultimate syllable of the clitic group, the syllabic structure being the same in both cases.

Moreover, a closer, comparative investigation into enclitic strings in Campidanese and in the Transitional area reveals that also in the latter variety the phonological form of clitic groups is only seemingly regular.

In both Campidanese and the Transitional area, the third person Accusative/Dative enclitics have a retroflex plosive geminate resulting from a lateral geminate, and therefore we take those forms as conservative outcomes of Latin ɪʟʟu.[3] In the same way, we take the form of the Locative/Partitive as the outcome of Latin ɪɴᴅᴇ (Wagner 1941). The examples in (8) show that in Northern Campidanese the outcomes of Latin ɪʟʟu and ɪɴᴅᴇ take slightly different forms, depending on the exact composition of the clitic string.

---

3. As in many other Romance languages, Accusative and Dative are syncretic in 3rd person clitics.

(8)   *Enclitic forms selected by different inflectional classes in Campidanese*[4]

		1SG/2SG and 1PL/2PL	locative and partitive clitics	3MSG/3FSG and 3PL.M/3PL.F
a.	*-ai*	'sɛɣa 'cut (Imp.)'	sɛɣa'mi	sɛ'ɣandi	sɛ'ɣaddụ sɛ'ɣaddạ
b.	*'-i(ri), -i(ri)*	'pɔni 'put(Imp.)'	pɔni'mi	pɔ'nɛndi	pɔ'neddụ pɔ'nɛddạ
		'betti 'bring (Imp.)'	betti'mi	bet'tɛndi	bet'teddụ bet'tɛddạ
		a'βeri 'open (Imp.)'	aβeri'mi	aβe'rɛndi	aβe'reddụ aβe'rɛddạ

(Tertenía)

In the forms in (8), a difference may be observed depending on the inflectional class of the verb, in that verbs in *-ai* (from Latin -ARE) are distinguished from verbs of other classes. While verbs in *-ai* preserve the vowel *-a-* before any kind of clitic (8a), the other verbs have *-i-* before first and second person singular and plural clitics, and a front mid vowel, which regularly undergoes metaphony, elsewhere (8b).

This vowel alternation strongly suggests that the third person Accusative/ Dative clitics and the Locative/Partitive clitic have stressed, disyllabic forms, that are conservative outcomes of Latin ILLU, INDE. We can think that the imperative ending *-i-* is preserved before the initial consonant of the *-mí, -dí, -sí* clitics (9a), while is deleted, by virtue of vowel elision, before the initial vowel of the ILLU, INDE clitics (9b). The same holds for the final vowel of the clitic *-endi,* wich is deleted before a following vowel (9c). Instead, the stressed *-i-* of the *-mí, -dí, -sí* is always preserved before a following clitic (9d):

(9)   a.   a'βeri – mi/ði/zi → aβeri'mi, aβeri'ði, aβeri'zi
      b.   a'βeri – ɛndi → aβe'rɛndi
           a'βeri – eddụ, ɛddạ → aβe'reddụ, aβe'rɛddạ
      c.   aβe'rɛndi – eddụ, ɛddạ → aβerin'deddụ, aβerin'dɛddạ
           aβeri'mindi – eddụ, ɛddạ → aβerimin'deddụ, aβerimin'dɛddạ
      d.   aβeri'mi – ɛndi → aβeri'mindi
           aβeri'mi – eddụ, ɛddạ → aβeri'middụ, aβeri'middạ   (Tertenía)

---

4. Logudorese has three conjugations: *-are, '-ere* (stress on preceding syllable) and *-ire.* In Campidanese the distinction between *'-ere* and *-ire* is preserved only in a different stress position in the infinitive, because of a phonological process that raised final mid vowels. Today, Campidanese has *-ai, '-i(ri* (stress on preceding syllable) *-i(ri).* See Jones (1997: 331) and Pisano (2007, 11), among others.

To sum up, verbs in *-ai* in Campidanese preserve the imperative ending *-a* before any clitic sequence, while in verbs of other classes the vocalic ending is deleted before clitics of the ɪʟʟᴜ, ɪɴᴅᴇ class.

Interestingly, a parallel investigation in the Transitional area paradigms reveals some discrepancies. In this variety, the imperative ending *-i* is retained in all kind of clitic sequence (10a). This behaviour is seemingly consistent with the analysis in (10b), whereby all clitic morphemes are monosyllabic and unstressed.

(10) a. [bat'timi] [bat'tiḍḍu] [bat'tiḍḍa] [bat'tiɳḍe] [battiɳ'ḍeḍḍu] [battiɳ'ḍɛḍḍa]
      b. /batti-mi/ /batti-ḍḍu/ /batti-ḍḍa//batti-nde/ /battinde-ḍḍu/ /battinde-
          ḍḍa/   (Villagrande Strisaili)

On the contrary, the disyllabic, stressed forms of third person and Locative/Partitive clitics can be easily detected in the forms of gerund, both in Campidanese and in the Transitional area. Given the gerund form in (11a), the clitic groups in (11b), (11c), (11d) are still compatible with an 'agglutinative analysis' like (10b). However, this is not the case, as forms in (11e), (11f) clearly show. When a third person or a Locative/Partitive clitic is attached to the simple gerund base, the disyllabic allomorphs emerge and the 'agglutinative analysis' (see 11e′, 11f′) is excluded. More specifically (see 11e″, 11f″), the sequences reveal the presence of the disyllabic ɪʟʟᴜ, ɪɴᴅᴇ clitics and of a dental consonant that prevent vocalic contact.

(11) *Gerund + clitic(s)*

		Transitional area	Campidanese
a.	bring-ᴘʀᴏɢ	bat'tiɳḍo	bet'tɛndu
b.	bring-ᴘʀᴏɢ = ᴅᴀᴛ.1sɢ	battiɳ'ḍomi	bettɛndu'mi
c.	bring-ᴘʀᴏɢ = ᴅᴀᴛ.1sɢ = ᴘᴀʀᴛ	battiɳḍo'minde	bettɛndu'mindi
d.	bring-ᴘʀᴏɢ = ᴅᴀᴛ.1sɢ = ᴘᴀʀᴛ = ᴀᴄᴄ.3sɢ.ᴍ	battiɳḍomiɳ'ḍeḍḍu	bettɛndumin'deḍḍu
e.	bring-ᴘʀᴏɢ = ᴀᴄᴄ.3sɢ.ᴍ	battiɳḍo'ðeḍḍu	bettɛndu'ðeḍḍu
e″.		*battiɳḍo-ḍḍu	*bettɛndu-ḍḍu
e′.		battiɳḍo-ð-eḍḍu	bettɛndu-ð-eḍḍu
f.	bring-ᴘʀᴏɢ = ᴘᴀʀᴛ	battiɳḍo'ðɛɳḍe	bettɛndu'ðɛndi
f′.		*battiɳḍo-ɳḍe	*bettɛndu-ndi
f″.		battiɳḍo-ð-ɛɳḍe	bettɛndu-ð-ɛndi
		(Villagrande Strisáili)	(Tertenía)

To sum up, if examined in detail, the phonological composition of clitic groups in Campidanese and the Transitional area shows that clitic adjunction is not a simple matter of morphemes agglutination, and instead implies the insertion of specialised morphs, specifically those that continue the Latin ɪʟʟᴜ, ɪɴᴅᴇ and that are disyllabic and stressed.

Although in Logudorese such allomorphy is not clearly detectable, data concerning enclitic stress in Sardinian on the whole do not support the idea that stress patterns in enlarged domains result from the application of the same rules holding for word stress. More generally, Sardinian data do not support the statement that stress shift phenomena have purely phonological motivations.

As far as the enclitic stress of the Transitional area is concerned, although the pattern is completely unproblematic from a metrical point of view, it cannot be accounted for as the result of a phonological pressure towards generalised trochaic foot. Such an account would undergo the same criticism we mentioned above about other Italo-Romance languages, i.e. that antepenultimate stress is allowed in lexical entries, as observed in Section 3.1. Therefore we argue that stress shift is not caused by metrical requirements, although it may optimally meet them.

The case of Campidanese is the most intriguing, since it shows variability also when a single enclitic is attached, e.g. *aberimí* 'open (sth) for me', *aberidí* 'open (sth) for yourself' vs. *aberéddu* 'open it (M)', *aberèndi* 'open some of those (PART)', (data from Tertenía). However, this peculiarity has a simple description if we take account of the mentioned allomorphic variation: in Campidanese, a single enclitic, whether it is monosyllabic or disyllabic, always takes stress, and stress always falls inside the clitic section of the group, not on the host.

It could be observed that the Campidanese pattern could also be accounted for by the 'weak pronoun' hypothesis, given that both the *-mí/ -dí/ -sí* and the *-eddu, -endi* series contain elements that are stressed and therefore, according to that approach, non-clitic. However, we will show in the next section that this is not the case.

5.   A note on proclitic stress in Sardinian

In this paragraph, we will briefly address the phonological properties of Clitic(s) + Verb sequences in Sardinian, in order to specifically test the 'weak pronoun' hypothesis on those data, independently of other, more general considerations about the approach by Ordóñez and Repetti (2006, 2014), already mentioned above. Unlike other Romance languages that have enclitic stress shift, Sardinian varieties do not show any difference in the phonological form of clitics between enclisis and proclisis (cf. Lai 2017: 193–196). The examples in (12) refer to Logudorese, those in (13) to the Transitional area, those in (14) to Campidanese. Due to space limitations, we only present glosses for the Logudorese examples in

(12). Since the examples in (13) and (14) are literal translations of the former, the same glosses apply.[5]

(12) *Proclisis vs Enclisis in Logudorese*
      vattire 'to bring'

mi 'vattiði	'vattimi
DAT.1SG = bring-PRES.IND.3SG	bring-IMP.2SG = DAT.1SG
'(S/he) brings (sth) to me'	'Bring (sth) to me!'
kɛ 'vattiði	'vattike
LOC = bring-PRES.IND.3SG	bring-IMP.2SG = LOC
'(S/he) brings (sth) there'	'Bring (sth) there!'
la 'vattiði, lu 'vattiði	'vattila, 'vattilu
ACC.3SG.F/M = bring-PRES.IND.3SG	bring-IMP.2SG = ACC.3SG.F/M
'(S/he) brings it (f/m)'	'Bring it (f/m)!'
'mila 'vattiði, 'milu 'vattiði	vatti'mila, vatti'milu
DAT.1SG = ACC.3SG.F/M = bring-PRES.IND.3SG	bring-IMP.2SG = DAT.1SG = ACC.3SG.F/M
'(S/he) brings it (f/m) to me'	'Bring it (f/m) to me!'
'kɛla 'vattiði, 'kelu 'vattiði	vatti'kɛla, vatti'kelu
LOC = ACC.3SG.F/M = bring-PRES.IND.3SG	IMP.2SG = LOC = ACC.3SG.F/M
'(S/he) brings it (f/m) there'	'Bring it (f/m) there!'
'mikɛ 'vattiði	vatti'mike
DAT.1SG = LOC = bring-PRES.IND.3SG	DAT.1SG = LOC = bring-IMP.2SG
'(S/he) brings (sth) there for me'	'Bring (sth) there for me!'
mi'kɛla 'vattiði, mi'kelu 'vattiði	vattimi'kɛla, vattimi'kelu
DAT.1SG = LOC = ACC.3SG.F/M = bring-PRES.IND.3SG	DAT.1SG = LOC = bring-IMP.2SG = ACC.3SG.F/M
'(S/he) brings it (f/m) there for me'	'Bring it (f/m) there for me!'

(data from Olièna)

(13) *Proclisis vs Enclisis in the Transitional area*
      attire 'to bring'

mi 'attiði	at'timi
ŋŋɛ 'attiði	at'tiŋŋe
ɖɖa 'attiði, ɖɖu 'attiði	at'tiɖɖa, at'tiɖɖu
'miɖɖa 'attiði, 'miɖɖu 'attiði	atti'miɖɖa, atti'miɖɖu
ŋ'ŋɛɖɖa 'attiði, ŋ'ŋeɖɖu 'attiði	attiŋ'ŋɛɖɖa, attiŋ'ŋeɖɖu
'miŋŋɛ 'attiði	atti'miŋŋe
miŋ'ŋɛɖɖa 'attiði, miŋ'ŋeɖɖu 'attiði	attimiŋ'ŋɛɖɖa, attimiŋ'ŋeɖɖu

(data from Tonára)

---

5. All the data in (12–14) are original and has been collected in Rosangela Lai's fieldwork.

(14)   *Proclisis vs Enclisis in Campidanese*
       bettiri 'to bring'
       mi 'ɛttidi   betti'mi
       ndi 'ɛttidi   bet'tɛndi
       ɖɖa 'ɛttidi, ɖɖu 'ɛttidi   bet'tɛɖɖa, bet'teɖɖu
       'miɖɖa 'ɛttidi, 'miɖɖu 'ɛttidi   betti'miɖɖa, betti'miɖɖu
       n'dɛɖɖa 'ɛttidi, n'deɖɖu 'ɛttidi   bettin'dɛɖɖa, bettin'deɖɖu
       'mindi 'ɛttidi   betti'mindi
       min'dɛɖɖa 'ɛttidi, min'deɖɖu 'ɛttidi   bettimin'dɛɖɖa, bettimin'deɖɖu
                                               (data from Tertenía)

Besides the complete parallelism of the forms in proclisis and enclisis, which is at
odds with the hypothesis that they belong to different categories, one more specific
aspect can be observed in all Sardinian varieties. As argued in Lai (2017: 194–196),
both enclitic and proclitic clusters receive stress when they attach to the host. This
is evidence that in Sardinian strings of clitics form an independent prosodic word
on their own, both in enclisis and in proclisis. The stress on proclitics, besides
being perceptible, is also revealed by the metaphonic alternation e/ɛ in the third
person Accusative/Dative clitics:

(15)   *Metaphony in enclisis and proclisis*
       min'dɛɖɖa 'ɛttidi, min'deɖɖu 'ɛttidi   bettimin'dɛɖɖa, bettimin'deɖɖu

Since only stressed vowels undergo metaphony, we can take those data as fur-
ther evidence that the same morphemes occur in both proclisis and enclisis in
Sardinian, and come to the conclusion that in these systems enclitic stress shift is
not due to the insertion of weak pronouns instead of true clitics.[6]

6.   Conclusions

The data we have examined in this paper concerning enclitic stress in Sardinian
contribute to a modelling of clitic stress shift phenomena in Romance languages.
Sardinian varieties differ from each other in significant ways in regard of enclitic
stress pattern, although they share the same parametric setting for lexical stress.
The picture that emerges from our examination does not support purely phono-
logical explanations, on the one hand, and cannot be represented in terms of a true
clitic/weak pronoun contrast, on the other hand. On the contrary, it reveals the
considerable role of allomorphy in the make-up of clitic groups. Allomorphy af-
fects, besides segmental content, the stress structure of the sequences, with results

---

6. For further discussion of metaphony and stress on clitic clusters, see Lai (2017).

that often go against the metrical preferences of these languages, whereby the optimal constituent is the trochaic foot.

This complex set of data can fall into place in the broader framework provided by Manzini & Savoia (2017) who take into account the variation concerning cliticisation that affects many Italian and Romance languages. This variation consists in reordering of the clitic string, alternation between presence and absence of /l/ in third person Accusative/Dative clitics (e.g. Lucanian /lə/ ~ /u/ 'him/it (м)') and in stress shift. Manzini & Savoia highlight that both l/Ø alternation and stress shift do not depend on phonological rules of general validity in the grammar; rather, they are due to allomorphy and are part of lexically stored information. Nevertheless, they argue, the distribution of allomorphs is governed by syntactico-semantic principles and principles of externalisation. In the authors' insight, stress shift in enclisis and the presence of *l*-allomorph correlate with 'non-veridical' contexts (e.g. presence of negation and V-in-C structures, i.e. imperative, infinitive clauses). Their hypothesis is that non-veridical contexts require the presence of forms that realise definiteness (D) properties, while non-D forms may be inserted elsewhere. In their explanation, clitic allomorphy strengthens the visibility of D properties in specific syntactic configurations by enhancing the phonological content of clitics. In this perspective, stressed enclitics, as well as *l*- insertion, contribute to the externalisation of abstract semantic information.

Manzini & Savoia's (2017) proposal provides a theoretical framework to analyses like the one here proposed, whose result is that the phenomena investigated cannot be attributed to general principles of prosodic parsing or syntactic feature checking. In this broader perspective, characterising certain alternations as allomorphic does not mean missing possible generalisations: rather, it contributes to define the role of lexicalised content in expressing syntactico-semantic meaning.

References

Bafile, Laura. 1993. *Fonologia prosodica e teoria metrica: accento, cliticizzazione e innalzamento vocalico in napoletano*. PhD dissertation, Università di Firenze.

Bafile, Laura. 1994. La riassegnazione postlessicale dell'accento nel napoletano. *Quaderni del Dipartimento di Linguistica dell'Università di Firenze* 5: 1–23.

Blasco Ferrer, Eduardo. 1984. *Storia linguistica della Sardegna*. Tübingen: Niemeyer. https://doi.org/10.1515/9783111329116

Blasco Ferrer, Eduardo. 1986. *La lingua sarda contemporanea. Grammatica del logudorese e del campidanese*. Cagliari: Edizioni Della Torre.

Blasco Ferrer, Eduardo. 1988. *Le parlate dell'Alta Ogliastra*. Cagliari: Edizioni Della Torre.

Bolognesi, Roberto. 1998. *The phonology of Campidanian Sardinian: A unitary account of a self-organizing structure* [HIL Dissertations 38]. The Hague: HAG.

Cardinaletti, Anna & Starke, Michal. 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.

Halle, Morris & Vergnaud, Jean-Roger. 1987. *An essay on stress* [Current Studies in Linguistics 15]. Cambridge MA: The MIT Press.

Halle, Morris. 1990. Respecting metrical structure. *Natural Language & Linguistic Theory* 8: 149–176. https://doi.org/10.1007/BF00208522

Halle, Morris & Kenstowicz, Michael. 1991. The free element condition and cyclic versus non-cyclic stress. *Linguistic Inquiry* 22(3): 457–501.

Jones, Michael Allan. 1997. Sardinian. In *Romance Languages*, Martin Harris & Nigel Vincent (eds), 314–350. London: Routledge.

Kenstowicz, Michael. 1991. Enclitic Accent: Latin, Macedonian, Italian, Polish. In *Certamen Phonologicum, II: Papers from the 1990 Cortona Phonology Meeting*, Pier Marco Bertinetto, Michael Kenstowicz & Michele Loporcaro (eds), 173–185. Torino: Rosenberg & Sellier.

Lai, Rosangela. 2016. Clitici e slittamento accentuale nelle varietà sarde: Tra prosodia e allomorfia. In *La lingua e i parlanti. Studi e ricerche di linguistica*, Benedetta Baldi & Leonardo M. Savoia (eds), 135–149. Alessandria: Edizioni Dell'Orso.

Lai, Rosangela. 2017. Stress shift under cliticization in Nuorese Sardinian. *Quaderni di Linguistica e Studi Orientali* 3: 183–199. http://dx.doi.org/10.13128/QULSO-2421-7220-21344

Manzini, M. Rita. 2014. Grammatical categories: Strong and weak pronouns in Romance. *Lingua* 150: 171–201. https://doi.org/10.1016/j.lingua.2014.07.001

Manzini, M. Rita & Savoia, Leonardo M. 2017. Enclisis/Proclisis alternations in Romance: Allomorphies and (re)ordering. *Transactions of the Philological Society* 115: 98–136. https://doi.org/10.1111/1467-968X.12093

Ordóñez, Francisco & Repetti, Lori. 2006. Stressed enclitics? In *New Analyses in Romance Linguistics* [Current Issues in Linguistic Theory 275], Jean-Pierre Montreuil & Chiyo Nishida (eds), 167–181. Amsterdam: John Benjamins.

Ordóñez, Francisco & Repetti, Lori. 2014. On the morphological restriction of hosting clitics in Italian and Sardinian dialects, *L'Italia Dialettale* 75: 173–199.

Peperkamp, Sharon. 1997. *Prosodic Words* [HIL Dissertations 34]. The Hague: HAG.

Pescarini, Diego. Forthcoming a. Stressed enclitics are not weak pronouns: A plea for allomorphy. In *Romance Linguistics 14, Selected papers from the 46th Linguistic Symposium on*

Romance Languages [Romance Languages and Linguistic Theory 14] Francisco Ordóñez & Lori Repetti (eds). *Amsterdam: Benjamins.*

Pescarini, Diego. forthcoming b. An emergentist view on functional classes. In *Linguistic Variation: Structure and Interpretation. Contributions in Honor of M. Rita Manzini* [Studies in Generative Grammar 132], Ludovico Franco et al. (eds). Berlin: Mouton de Gruyter.

Pisano, Simone. 2007. *Il sistema verbale del sardo moderno: Tra conservazione e innovazione*. Pisa: Edizioni ETS.

Pittau, Massimo. 1972. *Grammatica del sardo nuorese*. Bologna: Pàtron.

Steriade, Donca. 1988. Greek accent: A case for preserving structure. *Linguistic Inquiry* 19: 271–314.

Wagner, Max Leopold. 1984[1941]. *Fonetica storica del Sardo* [Historische Lautlehre des Sardischen]. Cagliari: Trois Editore.

Clitics and vowel epenthesis
A case study

Jacopo Garzonio and Silvia Rossi
University of Padova / Goethe University Frankfurt

Syncope and epenthesis have been treated as two closely related phenomena in traditional accounts: what syncope destroys, epenthesis restores. In this paper we present some cases of vowel epenthesis in the verbal domain in some Northern Emilian varieties where both syncope and epenthesis are rather restricted. It will be shown that apparently free alternations in some arhizotonic forms, like *lizì* vs. *alzì* 'you.PL read', are to be considered as the result of two different grammars. Only in one of these there is allomorphy of verb stems triggered by the interaction of morphosyntactic configurations, and morphophonological and phonotactic constraints.

Keywords: Emilian dialects, syncope, epenthetic vowels, syntax-phonology interface

1. Introduction

One of the most distinguishing features characterizing the development of Gallo-Italic varieties from Latin is the general loss of pre-tonic and post-tonic unstressed vowels (syncope), which gave rise to peculiar consonant clusters (cf. Lat. HOSPITALE(M) > Bolognese *zbde:l* 'hospital'; Lat. DENARIU(M) > Pied. *dnè* 'money'; Lat. PECCATU(M) > Lomb. *pcà* 'sin'; Rohlfs 1966: 137ff.; Loporcaro 2009: 106ff.). Closely related to this process of generalized syncope is the emergence of vowel epenthesis, i. e. the insertion of a vocalic segment, usually /a/, in order to syllabify otherwise unsyllabifiable consonant clusters (Rohlfs 1966: 471ff.; Savoia 1994).

Both processes are extremely evident in Emilian dialects, and have received a great deal of attention not only in the Italo-Romance dialectological literature, but also in more theoretically-oriented studies, as their treatment is rather problematic in many phonological theories (for a detailed review of this and for an alternative account see Passino 2013; See also Savoia 2015: 353ff.).

In the present paper, we focus on a specific dialectal area which shows mixed linguistic features between Lombard and Emilian. The data come from three varieties: Luzzara (Reggio Emilia), Rolo (Reggio Emilia) and Carpi (Modena). In this dialectal area syncope and vowel epenthesis are present but they are not as pervasive as in other Emilian dialects. Luzzara in particular has some other very clear Mantuan Lombard traits. We have chosen such area because we are investigating the following optional phenomenon as emerged from the fieldwork for the *Atlante Sintattico d'Italia* (ASIt, 'Syntactic Atlas of Italy'):[1]

(1) a. An [*al'zi*]   mai   di lebar.   (Luzzara)
        not read.2PL never of books

    b. An [*li'zei/li'zi/li'ze*] mai   di lebar.
        not read.2PL   never of books
        'You never read books.'

In (1) the stem of the verb in the 2nd plural form of the present indicative shows an alternation between *liz-*, without syncope and epenthesis, and *alz-*, where both phenomena occur. The alternation in (1) reminds systematic alternations, driven by different stress positions in verbal paradigms, that have already been pointed out by many authors (see among others Loporcaro (1998), Passino (2013) and references therein). A typical alternation is described by Passino (2013: 7–8), where the infinitival form and the singular present indicative forms behave differently because only the former is arhizotonic:

(2) a. LIGICARE > (lekkare) > [al'kar] [a lek]   (Ferrara)
        'to lick'   'I lick'

    b. NIVERE > (nivare) > [an've:r] [al 'najva]   (Bologna)
        'to snow'   'it snows'

However, as already pointed out, the alternation considered here is not systematic, but optional, in the sense that we have observed the two different variants in the same dialect. The aim of the following discussion is twofold: we intend to describe the variation regarding the syncope plus epenthesis alternation. Furthermore, we want to explain the factors involved in the observed variation. In particular, we will propose that in some varieties there are two competing grammars: in the first one etymological verb stems are preserved across the whole paradigm, while in the other there is an alternation between syncope plus epenthesis and etymological allomorphs.

---

1. The ASIt database is available at <http://asit.maldura.unipd.it/>. See also Benincà & Poletto (2007), and Di Nunzio, Garzonio & Pescarini (2014).

The paper is organized as follows: the dialectal area investigated and the methodology of the study are presented in Section 2, which also gives a sketchy outline of vowel epenthesis in Emilian dialects. Our research question is addressed and discussed in 3, while Section 4 concludes.

2.   Syncope and epenthesis in Luzzara, Carpi and Rolo

Data has been collected through a small questionnaire specifically designed to elicit the different forms discussed in (1). However, in order to better understand the extent of syncope and vowel epenthesis in the varieties, the questionnaire includes a short list of words, which are known from the literature to undergo both phenomena in dialects like Bolognese and Ferrarese. More specifically, we drew up a list of words like Lat. HOSPITALE(M) > Bol. *zbdɛ:l* 'hospital' and Lat. STOMACHUS > Bol. s*tamg*, culled from Loporcaro (1998) and Passino (2013), who in turn quote Coco (1970). Interestingly, nearly all of these words have syncopated forms in Bolognese, while, as we will see, the varieties investigated here are much less systematic.

Starting with Luzzara, we notice that syncope, and thus vowel epenthesis, is very sporadic and usually restricted to lexical items which display the phenomena in the whole Emilian area, like (*a*)*dman* 'tomorrow', *femna* 'female' and *smana* 'week'. Other lexical items which are syncopated in Bolognese lack the phenomenon in Luzzara and present instead similar developments as those of non-Emilian Gallo-Italic: *ospedal* 'hospital', *stomag* 'stomach', *buttoŋ* 'botton' (cf. Bol. *ptaŋ*) and *pegura* 'sheep' (cf. Bol. *pi:gra*).

Moving a few kilometers, Rolo and Carpi present syncope in more items, as for instance in *pegra* 'sheep' and Carpi also in *polvra* 'dust' (cf. Bol. *palvra* but *polver* in Rolo). As for other items, the forms are parallel to those for Luzzara.

Thus, from the point of view of the lexicon, it is possible to conclude that items with little or null paradigmatic variation (i.e. no stress movement) are quite stable and do not present alternations. However, in the verbal system, things are much more complicated, as we will show in the next section.

3.   Discussion

The ASIt questionnaires for Luzzara clearly show that verbal roots are not syncopated in rhizotonic forms, as expected.

(3) At= [ˈlezi]   e   [riˈlezi]   sempar al stes lebar
you= read.2s and re-read.2s always the same book
'You keep reading and re-reading the same book.'

However, as the example in (1) shows, repeated here as (4), in arhizotonic forms, speakers alternate between *liz-*, the regular etymological outcome, and *alz-*, a form with syncope and epenthesis.

(4) a.   An [*alˈzi*]   mai   di lebar.   (Luzzara)
not read.2PL never of books
b.   An [*liˈzei/liˈzi/liˈze*] mai   di lebar.
not read.2PL   never of books
'You never read books.'

A similar result has been observed for Carpi where, the 2nd plural present indicative form after preverbal negation alternates in a similar fashion.

(5) a.   An [*alzi:v*]   (Carpi)
not read.2PL
b.   An [*lezi:v*]
not read.2PL

Since the data we have for Luzzara show that syncope and epenthesis appear only when there is a proclitic (the preverbal negation *an* or a clitic pronoun with a VC form), a preliminary hypothesis is that the two phenomena are triggered by this item. In the grammar of speakers that produce (5a), since negation and VC clitics have a VC syllable structure, they combine with the initial C of verbs, forming a $VC_1 C_2 VC_3$ cluster which is not phonotactically acceptable and is resolved changing it into $VC_1 VC_2 C_3$. While this explanation can be applied to the phenomenon in Luzzara and Carpi, in Rolo it appears to be completely unrelated to phonotactics, since the syncopated and epenthetic form can be observed even in cases like the following ones, i.e. negative imperatives with postverbal negation, where there is no item in proclisis:

(6) a.   [*alˈzidi*] mia!   (Rolo)
read.2P not
'Don't read!'
b.   [*alˈkɛdi*] mia!
lick.2P not
'Don't lick!'

Given the high degree of dialectal (and in some cases even intra-speaker) variation, we assume that in varieties like Luzzara or Carpi, for speakers that produce (4a) and (5a), there are two allomorphes for the arhizotonic root (found for instance

in the 1st and 2nd plural forms, or the infinitive) of some verbs. In varieties like Rolo, only the variant with syncope and epenthesis is present. On the other hand, in dialects like Carpi and Luzzara the syncope plus epenthesis allomorph normally surfaces when it is required by phonotactics, for example to avoid /n l/ clusters. The relation between allomorphic alternations across paradigms and the phenomena considered here can be observed also in other varieties of this area, like Martignana (Cremona), where Manzini and Savoia (2005: III-343) observe alternations like the following.

(7) a. a ˈson kuntent da vˈdi:=t
          I am happy   of see.INF = you
          'I'm happy to see you.'

     b. a = l   vøi   ˈvødɐr
         I = him want see.INF
         'I want to see him.'

In (7a), as opposed to (7b) the presence of the object clitic requires the arhizotonic infinitival, which displays syncope.[2]

Furthermore, it seems that this type of distribution has supported the permanence of the combination of syncope plus epenthesis in the phonological system of varieties where it is otherwise very limited in the lexicon. From this point of view, the variety of Luzzara is particularly interesting because the presence of forms with syncope can be linked to a general restriction on the number of unstressed syllables in proclisis. This is clearly shown by the fact that clitic clusters in proclisis are tolerated only if they are monosyllabic, with a CVC form, while bisyllabic clusters of the VCVC type are excluded. In the collected data the preverbal negation is regularly dropped when there is either a V or a VC clitic before the verb. In (8) the clusters formed by 3rd person plural subject clitic, [i], or by 3rd person feminine singular subject clitic, [la], plus negation form a single syllable CVC and preverbal negation is preserved.[3] In (9), on the other hand, VC clitics like the reflexive/impersonal clitic [az], the 1st person object clitic [am], the 2nd person subject

---

2. A further example found in Manzini and Savoia (2005: I, 153) is San Benedetto Po (Mantua), where arhizotonic forms of the verb corresponding to 'to drink' display both syncope and metathesis: *a bevi* 'I drink', *at bevi* 'you(sing) drink', but *a vbʊm* 'we drink' and *a vbi* 'you(pl) drink'.

3. Luzzara presents the vocalic clitic [a] for the 1st person singular subject and for both the 1st and 2nd person plural subjects. We have not included these examples as there is no way to identify whether a sequence like [an] is a subject clitic plus negation cluster or the form of negation alone, which varies between [an] and [n] in this variety.

clitic [at] and the 3rd person masculine subject clitic [al] do not form a cluster and preverbal negation is absent.[4]

(8)   a.   CV clitic
           [*la ŋ*] = magna   mai   la früta   cla pütela.   (Luzzara)
            she = not = eats never the fruit that girl
           'That girl never eats fruits.'
      b.   V clitic
           [*i ŋ*]=   cumpra mai   dla früta li me sureli.
            they not= buy   never of.the fruit the my sisters
           'My sisters never buy fruit.'

(9)   a.   [az]=   dis   mia acsè.
            REFL= says not so
           'That's not how you say it.'
      b.   [*am*]= a   mia vest nisun.
            me= has not seen nobody
           'Nobody has seen me.'
      c.   [*at*]= cumpri mai   di pom.
            you= buy   never of apples
           'You never buy apples.'
      d.   L'è   Piero c [*al*]= völ   mia partir.
           it = is Peter that he= wants not leave
           'It's Peter who does not want to leave.'

We assume that this restriction on clitic clusters in proclisis is in fact an avoidance of feet with a final head in initial position (∪ ∪ –, or even ∪ ∪ ∪ –). This is in line with the general avoidance of long head-final feet in Northern Italo-Romance (Marotta 1999). The restriction is the reason for the presence of the VC_1C_2 allomorph (like *alz-*) in the cases considered here. In other words, we claim that epenthesis must co-occur with syncope in order to avoid forms like *an alizei*, with an epenthetic vowel (to avoid consonant clusters like [n l]) and the "etymological" root of the C_1VC_2 type. This means that, in accordance with Bafile (2003), this allomorphic alternation is further evidence that syncope and epenthesis are indeed related, even if epenthesis is not only a strategy to repair consonant clusters formed by syncope.

───────────────

4.   Notice that this type of strategy is similar to the first type of strategy described by Loporcaro (1998: 162) for the southern Emilian variety of Grizzana, where clitic clusters like /al t/ (feminine 3rd plural subject plus 2nd singular object) are realized as [a t]. However, Grizzana also has an alternative epenthetic strategy which yields [al ti].

More in general, there are two possible grammars in Luzzara: for some speakers the proclisis restriction combines with a phonotactic restriction on sequences like /n l/ in order to have a systematic *lez−/alz-* allomorphy throughout the verb paradigm. For others the phonotactic restriction can be violated, at least in some cases, and arhizotonic forms can display the "etymological" variant *liz-*, thus displaying a *lez−/liz-* allomorphy. This variation appears to be very dynamic since we have found both possibilities for the same speaker in Carpi.

4.   Conclusions

In this article, we have argued that epenthetic vowels found in syncope contexts in verb forms in varieties like Luzzara are part of an allomorphic alternation. This entails that viewing these items as part of the lexicon hinges on the more general question on whether allomorphy in verb paradigms is part of the lexicon. In general, it appears that this type of allomorphy is conditioned by two interacting aspects: the rhizotonic/arhizotonic alternation, which has a historical origin, and the presence of clitics, which is determined by syntax. Interestingly, the maintenance of syncope/epenthesis allomorphs appears to be related both to a phonotactic and a prosodic constraint.

Acknowledgments

References

Bafile, Laura. 2003. Syncope, epenthesis and syllable structure: The case of some Italian Dialects. *Rivista di Grammatica Generativa* 28: 19–29.

Coco, Francesco. 1970. *Il dialetto di Bologna: fonetica storica e analisi strutturale* Bologna: Forni.

Benincà, Paola & Poletto, Cecilia. 2007. The ASIS enterprise: A view on the construction of a syntactic atlas for the Northern Italian dialects. *Nordlyd* 34: 35–52.

Di Nunzio, Giorgio M., Garzonio, Jacopo & Diego Pescarini. 2014. ASIt: Atlante Sintattico d'Italia. A linked open data geolinguistic web application. In *Collaborative Research Practices and Shared Infrastructures for Humanities Computing*, Maristella Agosti & Francesca Tomasi (eds), 197–203. Padova: CLEUP.

Loporcaro, Michele. 1998. Syllable structure and sonority sequencing: Evidence from Emilian. In *Romance Linguistics: Theoretical Perspectives. Selected papers from the 27th Linguistic*

Symposium on Romance Languages (LSRL XXVII), Irvine, 20–22 February, 1997 [Current Issues in Linguistic Theory 160], Armin Schwegler, Bernard Tranel & Myriam Uribe-Etxebarria (eds), 155–170. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.160.13lop

Loporcaro, Michele. 2009. *Profilo linguistico dei dialetti italiani*. Bari: Laterza.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Marotta, Giovanna. 1999. 'Degenerate Feet' nella fonologia metrica dell'italiano. In *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI congresso della Società di Linguistica Italiana (Padova, 25–27 settembre 1997)*, Paola Benincà, Alberto Mioni & Laura Vanelli (eds), 97–116. Roma: Bulzoni.

Passino, Diana. 2013. A place in the lexicon for the epenthetic vowels of the Emilian dialects. *Lingue e Linguaggio* 12(1): 5–30.

Rohlfs, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti. Fonetica*. Torino: Einaudi.

Savoia, Leonardo M. 1994. Parametrical variation in prosodic licensing: Evidence from Italo-Albanian dialects. *Rivista di Grammatica Generativa* 20: 105–141.

Savoia, Leonardo M. 2015. *I dialetti italiani. Sistemi e processi fonologici nelle varietà di area italiana e romancia*. Pisa, Pacini.

Overabundance in Hungarian accusative pronouns

Maria Grossmann and Anna M. Thornton

University of L'Aquila

Overabundance is defined as the situation in which more than one inflected form is available to realize a single cell of an inflectional paradigm (Thornton 2011, 2012). Hungarian pronominal paradigms host several cases of overabundance in the accusative forms. Quantitative data from both diachronic and synchronic corpora are presented, showing that 1/2/3SG.ACC forms have an unbalanced distribution while 1PL.ACC and 2PL.ACC forms have a very balanced distribution, close to a 1:1 ratio between the two competing forms. One of the factors mentioned in the literature as favoring one of the two 1PL.ACC and 2PL.ACC forms, i.e., usage in emphatic / focus position, is illustrated and briefly discussed on the basis of corpus data.

Keywords: overabundance, Hungarian, pronouns, 1st person plural, 2nd person plural

1.   Overabundance

Overabundance holds when more than one inflected form is available to realize a single cell of an inflectional paradigm. A few examples are given in (1):

(1)   Italian   *sepolto / seppellito*   'bury.PST.PTCP'
      Spanish   *supiera / supiese*   'know.1/3SG.IMPF.SBJV'
      Czech   *jazyku / jazyka*   'language.GEN.SG'
      Hungarian *minket / bennünket* '1PL.ACC' (pronoun)

The phenomenon has recently come to the fore of discussion in theoretical morphology (Thornton 2011, 2012, to appear; Stump 2016: 147–151; Crysmann & Bonami 2016), and many examples from several languages have been observed and

described (cf. Thornton to appear for an overview).[1] The existence of overabundance poses a theoretical problem for most models of morphology, which include principles (such as Blocking, or the Elsewhere Condition or Pāṇini's principle) by which a choice is always forced between two semantically equivalent forms.

Some authors deny even the theoretical possibility of overabundance: they maintain that even if there is no apparent stylistic or grammatical difference in the contexts in which the competing forms are used, such a difference must exist, and it is only our inability to find a way to detect it that obscures it (Berruto 1995: 159, fn. 2, is a champion of such a position). Kypriotaki (1974: 343) rightly observes that such a position is never falsifiable.

Other authors (e.g. Kroch 1994) acknowledge the existence of overabundance, but maintain that it is always diachronically unstable: one of the competing forms "tends to drive the other out of use and so out of the language" (Kroch 1994: 17).

Aronoff (2016, to appear) has recently observed that competition between two synonymous linguistic elements (lexemes, lexeme-formation rules, or inflectional realizations) is an instance of a more general property of biological systems, in which the principle of competitive exclusion holds. This principle states that no two species with similar ecological niches can coexist in a stable equilibrium. Aronoff draws a parallel between biological species and linguistic elements: as species compete for resources and habitats, so do linguistic elements. Resources and habitats in languages are represented by the environments in which a given element can occur, and the elements with which it can combine. In biology, when there is competition between two species for the same resources, two outcomes are possible: either one species wins and the other becomes extinct, or both species manage to survive thanks to adaptive changes in one or both of them. While Kroch's (1994) position, with respect to the competition between linguistic elements, acknowledges only the first possibility, Aronoff (2016, to appear; Lindsay & Aronoff 2013; Aronoff & Lindsay 2014) argues that competition between linguistic elements need not be resolved through the extinction of one of the competitors if each of the competing forms finds a niche in which it is preferably used. A niche is "a clearly defined subdomain" (Lindsay & Aronoff 2013: 135) within the domain in which the two competing elements can occur. Aronoff has studied several cases of competition, mainly between derivational affixes that have the same input and

---

1. An anonymous reviewer questions the usage of the term *overabundance*, and suggests using the general term *optionality*. We choose to stick to *overabundance*, defined as in the first sentence of the present paper. We consider it useful to have a term that refers specifically to the availability of different forms to realize a cell in an inflectional paradigm, distinct from the terms used to refer to other kinds of competing forms (e.g., different derivatives from the same base, or different synonymic syntactic constructions). For further observations on the terminology, cf. Thornton (2011: 359–360, 2016).

output category and semantic function (such as English *-ity* and *-ness*, *-ize* and *-ify*), and has shown how the less productive suffix has managed to survive by selecting a specific kind of bases, a phonologically, morphologically, semantically or pragmatically defined subdomain within the general domain of the two synonymous lexeme-formation rules. Aronoff also acknowledges that finding a niche is more difficult for "inflectional doublets", i.e. inflectional forms in a relation of overabundance: in this case, semantic differentiation is not possible because the meaning of a cell in an inflectional paradigm is defined by an intersection of syntactic feature-values.

However, putative cases of overabundance of inflectional realizations for a single paradigmatic cell keep coming up.

Recently, the advent of computational and statistical methods for linguistic analysis has allowed to investigate in detail cases of competition, both in syntax (e.g., Bresnan et al. 2007) and morphology (e.g., Bermel & Knittl 2012); the existence of stable cases of overabundance, with true interchangeability of forms and constructions in the same context, has been confirmed. In many cases, preferential constraints (Thuilier 2012) have been uncovered, but the competing forms or constructions have not been found to be in fully complementary distribution; significant overlapping between the contexts of occurrence of the two elements has been shown to occur.

Several cases in point occur in Hungarian, to which we now turn.

2.   Hungarian nominal and pronominal inflection: An overview

Hungarian nouns inflect for case, number and possessor's person.[2] Some examples of singular forms from the paradigm of кέz 'hand' are shown in Table 1.

Most case forms of Hungarian personal pronouns are built by inflecting for the appropriate person and number values dedicated stems that are homophonous to nominal case endings (or are partially or totally suppletive allomorphs of these endings). Like case endings, these stems originate from postpositions, which in turn originate from nouns (cf. D. Mátai 2003; Creissels 2006; Spencer 2008; and Spencer & Stump 2013 for a detailed presentation and analysis).

---

2. Hungarian displays vowel harmony; we don't discuss it here because it plays no role in the phenomena analyzed in this paper. When a linking vowel subject to vowel harmony appears between base and affix or between two affixes, we will indicate this vowel as V. We will also not discuss an important syntactic property: first and second person pronouns select the indefinite conjugation, while third person pronouns select the definite (cf., among others, Bartos 2000; Szili 2000; den Dikken 2004, 2006; Bárány 2015).

Table 1. Examples of forms of the noun κέz 'hand'

Case / Possessor	UNPOS- SESSED	1SG	2SG	3SG	1PL	2PL	3PL
NOMINATIVE	kéz	kez-em	kez-ed	kez-e	kez-ünk	kez-etek	kez-ük
INESSIVE	kéz-ben	kez-em-ben	kez-ed-ben	kez-é-ben	kez-ünk-ben	kez-etek-ben	kez-ük-ben
ABLATIVE	kéz-től	kez-em-től	kez-ed-től	kez-é-től	kez-ünk-től	kez-etek-től	kez-ük-től
ACCUSATIVE	kez-et	kez-em-et	kez-ed-et	kez-é-t	kez-ünk-et	kez-etek-et	kez-ük-et

The person / number exponents are the same ones used to inflect nouns for possessors, as shown in Table 2 for nominative, inessive and ablative forms.[3]

Table 2. Examples of personal pronouns inflected for case

Case / Person	1SG	2SG	3SG	1PL	2PL	3PL
NOMINATIVE	én	te	ő	mi	ti	ők
INESSIVE	benn-em	benn-ed	benn-e	benn-ünk	benn-etek	benn-ük
ABLATIVE	től-em	től-ed	től-e	től-ünk	től-etek	től-ük

Accusative forms are peculiar in several ways. They are shown in Table 3.

Table 3. NOMINATIVE and ACCUSATIVE forms of personal pronouns.

Case / Person	1SG	2SG	3SG	1PL	2PL	3PL
NOMINATIVE	én	te	ő	mi	ti	ő-k
ACCUSATIVE	eng-em	tég-ed	ő-t	mi-nk-et	ti-tek-et	ő-k-et
	eng-em-et	tég-ed-et	ő-t-et	benn-ünk-et	benn-etek-et	

It is apparent from the data in Table 3 that accusative forms of pronouns display overabundance in almost all persons.[4] Overabundance is of two kinds, which will be discussed separately: 1) unbalanced overabundance in 1/2/3SG.ACC forms; 2) balanced overabundance in 1PL.ACC and 2PL.ACC forms. In this paper, we will discuss the second kind more thoroughly.

---

3. The same exponents are used also in postpositions. For example, see the inflection of *mellett* 'beside': *mellett-em* 'beside me', *mellett-ed* 'beside you.SG', *mellett-e* 'beside him/her', *mellett-ünk* 'beside us', *mellett-etek* 'beside you.PL', *mellett-ük* 'beside them'.

4. In MNSZ2 there is also one token of substandard 3PL.ACC *őketet* (*ő-k-et-et* '3-PL-ACC-ACC').

3. Overabundance in Hungarian personal pronouns

3.1 Unbalanced overabundance in 1/2/3SG.ACC forms

The 1SG.ACC and 2SG.ACC forms *engem* and *téged* were formed by adding an ancient pronominal formative *-g* and the possessive markers to the nominative form (cf. D. Mátai 2003: 214); according to various scholars, they are open to various synchronic analyses, which we won't discuss here (cf. Thornton to appear for an overview, based on Rounds 2009: 114; Spencer 2008: 48; Kenesei, Vago & Fenyvesi 1998: 270); the 3PL forms bear the regular plural marker *-(V)k*; the marker *-(V)t* in 3SG.ACC *őt* and 3PL.ACC *őket* is the regular accusative ending.

The first kind of overabundance arises because the accusative ending *-(V)t* is optionally realized in 1/2SG.ACC forms and optionally reduplicated in 3SG.ACC form. This ending is optional also in possessed 1/2SG accusative forms of nouns and some other pronouns.[5] In 3SG.ACC *őtet* (*ő-t-et* '3-ACC-ACC') the accusative marker is reduplicated; this is usually explained by invoking a tendency to strengthen and lengthen the short form *őt*.[6]

The forms *engem / engemet, téged / tégedet, őt / őtet* have coexisted since the Old Hungarian period,[7] but their conditions of occurrence have changed in the course of history: in the past the forms displaying *-et* were more frequent and were used in more formal registers and in emphatic function; nowadays, they are less frequent and (particularly *őtet*) considered substandard from the diatopic and diastratic point of view (cf. D. Mátai 2003: 214, 399, 636, 740–741, 824).

---

5. For example,

  (i) *megmosom a   kez-em / kez-em-et*
      wash   the hand-1SG.POSS / hand-1SG.POSS-ACC
      'I wash my hands'

  (ii) *szeresd mag-ad / mag-ad-at*
       love   self-2SG.POSS / self-2SG.POSS-ACC
       'love yourself'
       (cf. Bartos 2000; Kiefer 2000; den Dikken 2004, 2006).

6. A similar phenomenon appears in the accusative forms of the demonstratives *ez* 'this' and *az* 'that', which, besides the regular accusative forms *ez-t* 'this-ACC' and *az-t* 'that-ACC', have substandard low frequency forms with a reduplicated ACC marker: *ez-t-et* 'this-ACC-ACC', *az-t-at* 'that-ACC-ACC'. Even more substandard are the forms *ez-t-et-et* 'this-ACC-ACC-ACC' and *az-t-at-at* 'that-ACC-ACC-ACC', which have extremely low frequency in our corpora.

7. The following periods are recognized in the history of Hungarian: Ancient Hungarian (Proto-Hungarian) 1000 BCE – 896 CE, Old Hungarian 896–1526, Middle Hungarian 1526–1772, Modern Hungarian 1772–1920, Contemporary Hungarian 1920-.

Table 4. Frequency of 1/2/3sG.ACC forms in the OHC, HDC and MNSZ2 corpora

Form1	OHC	HDC	MNSZ2	Form1 + −*et*	OHC	HDC	MNSZ2	Ratio in OHC	Ratio in HDC	Ratio in MNSZ2
1sG.ACC *engem*	243	8237	114170	*engemet*	673	1628	3573	1:2.8	5.1:1	32:1
2sG.ACC *téged*	326	2903	44843	*tégedet*	322	563	773	1:1	5.2:1	58:1
3sG.ACC *őt*	249	11230	156598	*őtet*	2085	1021	786	1:8.4	11:1	199.2:1

The frequency of these forms in Contemporary Hungarian and in older stages of the language is shown in Table 4 by means of corpus data.[8]

The ratio between the frequency of the two types has changed in time. In Old Hungarian *engemet* and *őtet* are more frequent than their competitors, while *téged* and *tégedet* are equally frequent. The frequency ratios change direction early in Modern Hungarian, making *engem*, *téged*, and particularly *őt* more frequent than their competitors by the end of the Modern Hungarian period; Table 5 illustrates the changes in the ratios between the frequency of the two forms in the course of the Modern Hungarian period.

Table 5. Ratio between the frequency of forms without and with *-et* in Modern Hungarian (data from HDC)

Time period Forms	1772– 1799	1800– 1819	1820– 1839	1840– 1859	1860– 1879	1880– 1899	1900– 1920
engem: *engemet* '1SG.ACC'	0.8:1	1.6:1	2.7:1	4.4:1	2.9:1	5.4:1	8:1
téged: *tégedet* '2SG.ACC'	1.3:1	1.4:1	2.8:1	4.5:1	4.7:1	5.5:1	10.4:1
őt: *őtet* '3SG.ACC'	0.4:1	0.2:1	3.3:1	21.3:1	16.2:1	56.7:1	18.2:1

In Contemporary Hungarian as represented both in HDC and in MNSZ2 the decrease in frequency of *engemet*, *tégedet* and *őtet* continues; in MNSZ2 *engemet* and *tégedet* are less frequent than their competitors by one order of magnitude, and *őtet* by two orders of magnitude.

---

8. We have extracted the frequency of the forms from several corpora, made available by the Research Institute for Linguistics (Hungarian Academy of Sciences):

1. Hungarian Gigaword Corpus (henceforth, MNSZ2), an extended new edition of the Hungarian National Corpus containing about 1.2 billion tokens. This corpus of present-day Hungarian contains texts from Hungary and neighboring countries where Hungarian is spoken, and can be queried by geographical areas and by text genres. Genres are the following: Journalism (daily/weekly newspapers), Literature (Digital Literary Academy), (Popular) science (Hungarian Electronic Library), Personal (social media), Official (documents from public administration), (Transcribed) spoken (radio programs).

2. Historical Dictionary Corpus (henceforth, HDC), containing about 17 M tokens, from mostly literary texts published between 1772 and 1994; it can be queried by year / time period and by text genres (prose, poetry, theatre).

3. Old Hungarian Corpus (henceforth, OHC), containing all extant records from the Old Hungarian period (896–1526), as well as some records from the Middle Hungarian period (1526–1772), totaling 2.2 M tokens.

Thornton (2012) has proposed to take the ratio between the frequency of two forms realizing the same cell as an indication of the strength and character of overabundance in that cell. A corollary of the definition of canonical overabundance is that two forms realizing the same cell should have approximately the same frequency of occurrence: if there is no factor conditioning the selection of either form, the two should be selected randomly, and have an equal chance of appearing in any context; this would lead to roughly equal frequency of the two forms in a corpus. Unbalanced frequencies indicate that the less frequent form is subject to conditioning, or declining in usage; this appears to be the case with Hungarian 1/2/3SG.ACC forms *engemet*, *tégedet*, *őtet*, which have become substandard.

3.2   Balanced overabundance in 1PL.ACC and 2PL.ACC forms

The second kind of overabundance is displayed by 1PL.ACC and 2PL.ACC forms.

The morphology of these forms is partially different from the previously observed ones. The forms *minket* and *titeket* (henceforth, pronominal stem forms) can be analyzed as in (2), while *bennünket* and *benneteket* (henceforth, *benn-* stem forms) are analyzed as in (3):

(2)   a.   *mi- nk-   et*
            1PL 1PL.POSS ACC
      b.   *ti- tek-   et*
            2PL 2PL.POSS ACC

(3)   a.   *benn- ünk-   et*
            benn 1PL.POSS ACC
      b.   *benn- etek-   et*
            benn 2PL.POSS ACC

The stem glossed 'benn' in (3) has the same origin as the inessive case marker (see Table 2); both derive from a postposition, itself originating from a noun, BÉL 'internal part', hence later 'bowels', followed by the locative ending *-n,* with assimilation of /ln/ to /nn/ (cf. Zaicz 2006, s.v. *benn*). D. Mátai (2003: 213, 399) shows that in Old Hungarian, besides *minket* and *titeket* (and its variant *ti-k-tek-et* '2PL-PL-2PL.POSS-ACC'), initially only the forms *bennünk* and *bennetek*, lacking an accusative marker, were found, with inessive meaning (these forms still exist as inessive forms, cf. Table 2); they subsequently developed a partitive accusative meaning ('some_of_us.ACC'), and at the end of the Old Hungarian period they already occur with the ACC marker *-et*, like today. In the OHC there are 246 tokens of *minket*, 288 tokens of *titeket*, no tokens of *bennünk*(*et*) and *benneteket*, and 2

tokens of *bennetek* used as an object.[9] According to D. Mátai (2003: 399, 636), in Old and Middle Hungarian *minket* and *titeket* are more frequent than *bennünket* and *benneteket*.

Rounds (2009: 114) states that nowadays the two forms of 1PL.ACC and 2PL. ACC pronouns "are used in free variation with no difference in style or meaning". If this assessment is correct, Hungarian 1PL.ACC and 2PL.ACC would constitute a rare case of canonical overabundance, i.e., a case in which the selection of one or the other form is subject to no conditions, either grammatical or geo-socio-stylistic (cf. Thornton to appear for a typology of conditions on overabundance, and many examples).

The frequency data, shown in Tables 6–14, show that the ratio between the two 1PL.ACC and particularly the two 2PL.ACC forms has been remarkably close to the 1:1 ratio of canonical overabundance since the end of the 18th century, and in all geographical areas and text genres.[10]

Table 6. Overall frequency of 1PL.ACC and 2PL.ACC forms in HDC and MNSZ2

Pronominal stem	HDC	MNSZ2	*benn-* stem	HDC	MNSZ2	Ratio HDC	Ratio MNSZ2
minket	2199	80511	*bennünket*	3667	50560	0.6:1	1.6:1
titeket	322	15918	*benneteket*	550	15026	0.6:1	1.1:1

Table 7. Frequency of 1PL.ACC pronouns in MNSZ2, by geographical area

	minket	*bennünket*	ratio
Hungary	78234	46497	1.7:1
Transylvania (Romania)	461	273	1.7:1
Subcarpathia (Ukraine)	215	323	0.7:1
Slovakia	1487	2937	0.5:1
Vojvodina (Serbia)	114	530	0.2:1

---

9. The first token is in the *Halotti beszéd és könyörgés* [Funeral Sermon and Prayer], the oldest known and surviving contiguous Hungarian text dating from 1192–1195: *ilezie wt Eſ tiv bennetuc* 'may [the Lord] resuscitate him/her and you.PL'; here *bennetuc* is preceded by the 2PL. NOM form *tiv*, for emphasis; the second token is in the *Jordánszky-Kódex* [Codex Jordánszky], a copy of the Bible's translation by László Báthory dating from 1516–1519: *boczaat … ennehanyat bennetek foghſagra* 'sends … some of you.PL in jail'.

10. For the HDC, the total number of tokens of each form in the different Tables is not always the same: for example, the total frequency of *minket* when queried in the whole corpus is 2199, when queried by genre is 2242, when queried by time period is 2408. We have no explanation for these differences, which are usually quite small.

Table 8. Frequency of 1PL.ACC pronouns in MNSZ2, by genre

	minket	*bennünket*	ratio
Official	3886	5668	0.7:1
Journalism	17384	13635	1.3:1
(Transcribed) spoken	8155	7147	1.1:1
Personal	43929	15401	2.9:1
Literature	5143	5617	0.9:1
(Popular) science	2014	3092	0.7:1

Table 9. Frequency of 2PL.ACC pronouns in MNSZ2, by geographical area

	titeket	*benneteket*	ratio
Hungary	15718	14762	1.1:1
Transylvania (Romania)	66	58	1.1:1
Subcarpathia (Ukraine)	23	19	1.2:1
Slovakia	106	170	0.6:1
Vojvodina (Serbia)	5	17	0.3:1

Table 10. Frequency of 2PL.ACC pronouns in MNSZ2, by genre

	titeket	*benneteket*	ratio
Official	41	79	0.5:1
Journalism	634	653	1:1
(Transcribed) spoken	162	236	0.7:1
Personal	14408	12989	1.1:1
Literature	479	811	0.6:1
(Popular) science	194	258	0.8:1

Table 11. Frequency of 1PL.ACC pronouns in HDC by genre

	Pronominal stem		***benn-* stem**		ratio
Prose	*minket*	1695	*bennünket*	3320	0.5:1
Poetry	*minket*	346	*bennünket*	177	2:1
Theatre	*minket*	201	*bennünket*	201	1:1

Table 12. Frequency of 2PL.ACC pronouns in HDC by genre

	Pronominal stem		benn- stem		ratio
Prose	*titeket*	201	*benneteket*	432	0.5:1
Poetry	*titeket*	84	*benneteket*	44	1.9:1
Theatre	*titeket*	42	*benneteket*	77	0.5:1

Table 13. Frequency of 1PL.ACC pronouns in HDC by time period

	Pronominal stem		benn- stem		ratio
1772–1799	*minket*	224	*bennünket*	270	0.8:1
1800–1849	*minket*	283	*bennünket*	429	0.7:1
1850–1899	*minket*	330	*bennünket*	446	0.7:1
1900–1949	*minket*	590	*bennünket*	1174	0.5:1
1950–1969	*minket*	260	*bennünket*	499	0.5:1
1970–1994	*minket*	521	*bennünket*	855	0.6:1

Table 14. Frequency of 2PL.ACC pronouns in HDC by time period

	Pronominal stem		benn- stem		ratio
1772–1799	*titeket*	54	*benneteket*	69	0.8:1
1800–1849	*titeket*	90	*benneteket*	87	1:1
1850–1899	*titeket*	50	*benneteket*	43	1.2:1
1900–1949	*titeket*	57	*benneteket*	182	0.3:1
1950–1969	*titeket*	33	*benneteket*	86	0.4:1
1970–1994	*titeket*	38	*benneteket*	84	0.5:1

4.   Are there differences in the usage of the two competing forms?

In and of themselves, balanced ratios do not prove that two forms that realize the same cell are used interchangeably: they could even be in complementary distribution, with each one appearing in a similar number of environments, disjunct from the ones in which the other form appears. Therefore, we must proceed, in the little space that remains, to investigate whether there are conditions on the selection of any one of the forms.

It is certainly not the case that the two 1PL.ACC and 2PL.ACC forms are in complementary distribution. They occur in comparable contexts, such as the ones in (4) and (5):[11]

(4) a. *Minket elsősorban az foglalkoztatott, hogy …*
   us   first_of_all that concerned   that
   'We were primarily concerned by…'

b. *Bennünket elsősorban az érdekelt, hogy …*
   us   first_of_all that interested that
   'We were primarily interested in…'

(5) a. *Minket is meglepett az eredmény*
   us   too surprised the result
   'The result surprised us too'

b. *Bennünket is meglepett a szavazás végeredménye*
   us   too surprised the voting   final_result
   'The final result of the voting surprised us too.'

---

11. The question is often raised whether two forms in an overabundant cell also occur in speech or writing by one and the same speaker. Data from other case studies show that it is possible to find two competing forms used by the same speaker (e.g., cf. Cappellaro 2013: 213 for Italian *sepolta / seppellita* 'buried.F.SG'). In the Hungarian corpora, we find many examples in which two different 1PL or 2PL pronouns are used by the same speaker/writer even within the same sentence; examples (i) and (ii) are from MNSZ2 (abbreviations: CAUS = causal-final, SUBL = sublative, SUPERESS = superessive, TRANSL = translative):

(i) *A magyar nyelv és a kultúra az, amely*
   DEF Hungarian language and DEF culture that which
   összetart **minket**, *ez tesz nemzetté* **bennünket**
   hold_together **1PL.ACC** this make nation.TRANSL **1PL.ACC**
   'The Hungarian language and culture, that's what keeps **us** together, what makes **us** a nation'

(ii) *Szeretünk* *Titeket ezért úgy döntöttünk,*
   love.PRS.1PL.INDF **2PL.ACC** this.CAUS so decide.PST.1PL.INDF
   nem hagyunk *benneteket koncert nélkül péntek estére*
   not leave. PRS.1PL.INDF **2PL.ACC** concert without Friday evening.SUBL
   Budapesten!
   Budapest.SUPERESS
   'We love **you**, that's why we decided not to leave **you** without a concert on Friday night in Budapest!'

The literature offers only short and sporadic observations on differences in the usage of *minket*, *titeket* vs. *bennünket*, *benneteket*.[12] In the following section we offer some preliminary observations based on corpus data concerning one of the factors that have been considered relevant in the previous literature.[13]

4.1 Emphatic vs. non-emphatic

Several grammars and dictionaries observe that *minket* and *titeket,* as opposed to *bennünket* and *benneteket,* are used to emphasize their referents (cf. also D. Mátai 2003: 741 with respect to Modern Hungarian). However, we do not know of any corpus-based study devoted to this issue. We offer here some preliminary data from the MNSZ2 that bear on the issue.

In Hungarian, emphasis "is expressed by means of varying the constituent order and assigning stress to some distinguished element" and "is usually discussed under the term 'focus'" (Kenesei, Vago & Fenyvesi 1998: 161). The emphatic or focus position is the immediately pre-verbal one (Kenesei, Vago & Fenyvesi 1998: 163; É. Kiss 2002: 77ff.). Therefore, investigating the occurrence of the 1PL. ACC and 2PL.ACC pronouns in pre-verbal vs. post-verbal position could give an indication of whether any form is preferred in emphatic / focus position.

Table 15 presents data extracted from the MNSZ2 corpus for a sample of 17 verbs.[14]

---

12. We must mention at least the short articles by Hunfalvy (1856), Steiner (1873) and Tálos (1980). Hunfalvy (1856) maintains that *bennünket* and *benneteket* have a partitive meaning ('part of us/you', vs. *minket / titeket* 'us/you all'). According to Steiner (1873)*bennünket* and *benneteket* are used to refer to a lesser quantity of people and are non-emphatic forms, while *minket* and *titeket* are emphatic. Tálos (1980), while maintaining that the two pairs are in free variation, and that any single form is selected for stylistic purposes, suggests that in some cases their distribution can be explained by appealing to the categories 'inclusive' (*bennünket* and *benneteket*) and 'exclusive' (*minket* and *titeket*). According to Tálos the exclusive forms are the unmarked ones, and can also occur with inclusive function, particularly in topic position, while the use of inclusive forms with exclusive meaning, while not leading to misunderstandings, sounds inappropriate. To explain the inclusive meaning of *bennünket* and *benneteket* Tálos reminds us of the reinterpretation inessive > partitive inessive > partitive accusative > accusative which took place in the history of the two forms. Corpus-based investigation of these factors must be left for further research, for lack of space.

13. Many more factors should be considered; we leave these for future research, for space considerations.

14. The verbs, chosen among the semantically most basic ones, are the following ones: *akar* 'want', *gyűlöl* 'hate', *hall* 'hear', *hallgat* 'listen', *hív* 'call, summon', *hoz* 'bring', *imád* 'adore', *kér* 'ask, request', *követ* 'follow', *lát* 'see', *néz* 'look', *szeret* 'love, like', *támad* 'attack', *utál* 'abhor, detest', *vár* 'wait', *véd* 'defend', *visz* 'carry, take to a place'.

Table 15. Frequency of 1PL.ACC and 2PL.ACC pronouns in immediately pre- and post-verbal position (on a sample of 17 verbs)

Context	*minket*	*bennünket*	ratio	*titeket*	*benneteket*	ratio
Pronoun V	1376	461	3:1	223	61	3.7:1
V Pronoun	8963	4251	2.1:1	5387	5410	1:1
Total	10339	4712	2.2:1	5610	5471	1:1

The data show that indeed *minket* and *titeket* are favored in pre-verbal, i.e. emphatic / focus position: the ratio of pronominal stem forms to *benn-* forms is 3:1 or above in this position, while it is more balanced in the post-verbal position (1:1 for the 2PL.ACC forms).

Other data bearing on the issue of whether the pronominal stems are favored in emphatic / focus position come from considering the usage of the 1PL.ACC and 2PL.ACC pronouns when following *csak* 'only'. This "particle" (as it is called in the literature on Hungarian) "generally occurs in the preverbal focus position" (Kenesei, Vago & Fenyvesi 1998: 164) and according to É. Kiss (2002: 104) bears an inherent [+ focus] feature, that it lends to its complement. Table 16 illustrates the frequency of the 1PL.ACC and 2PL.ACC pronouns when following *csak*: in this context, *minket* and *titeket* are clearly favored over their *benn-* stem counterparts, with ratios above 7:1.

Table 16. Frequency of the 1PL.ACC and 2PL.ACC pronouns when following *csak* 'only'

Context	*minket*	*bennünket*	ratio	*titeket*	*benneteket*	ratio
csak + Pronoun	447	62	7.2:1	38	5	7.6:1

5. Discussion and concluding remarks

In the SG.ACC pronouns we have geo-socio-stylistically conditioned overabundance: the forms *engemet*, *tégedet*, *őtet,* which started out as the more frequent forms, have declined in usage and nowadays they are not recommended by the norm and perceived as substandard. However, it is notable that both forms in each of the SG cells of Hungarian personal pronouns have been attested all along the history of the Hungarian language for which we have corpus documentation. This is a remarkable case of maintenance of overabundance in diachrony.

There do not seem to be strong geo-socio-stylistic conditions at work in the choice of the pronominal vs. *benn-* form of the 1PL.ACC and 2PL.ACC pronouns. They are certainly used with different frequency in different time periods,

geographical areas and text genres (cf. Tables 7–14), but both forms of both pronouns are amply represented in all areas, genres and time-periods.

The literature suggests that for 1PL.ACC and 2PL.ACC pronouns there might be grammatical conditions at play, among which we have briefly investigated the role played by emphatic function. There certainly appears to be a preference for the pronominal stem forms *minket* and *titeket* when the accusative pronoun is in focus / emphatic position; however, corpus data show that the *benn-* stem forms are by no means excluded in such positions. Many contexts can be found in which the two competing forms of each pronoun appear to be interchangeable. The case of Hungarian 1PL.ACC and 2PL.ACC pronouns seems to be a case very close to canonical overabundance, in which conditions on the choice of one or the other form have given rise to preferences, but not (yet?) to niches in which one of the two competing forms is entrenched. And, as in the case of SG.ACC pronouns, overabundance in these cells has been attested for centuries, and in this case with a very balanced ratio between the frequency of the two forms. The situation does not appear to be diachronically unstable, and none of the competing forms seems to be at risk of being driven out of usage; preferences for the usage of certain forms in certain contexts exist, but many contexts in which the two forms are interchangeable also exist; the corpora give evidence that the two forms can even be used by the same speaker within the same utterance. The Hungarian 1PL.ACC and 2PL.ACC pronouns represent one of the most canonical cases of overabundance discovered so far.

Acknowledgments

References

Aronoff, Mark. 2016. Competition and the lexicon. In *Livelli di analisi e fenomeni di interfaccia. Atti del XLVII congresso internazionale della Società di Linguistica Italiana*, Annibale Elia, Claudio Iacobini & Miriam Voghera (eds), 39–52. Roma: Bulzoni.

Aronoff, Mark. To appear. Competitors and alternants. In *Competition in Morphology*, Franz Rainer, Francesco Gardani, Hans-Christian Luschützky & Wolfgang U. Dressler (eds). Dordrecht: Springer.

Aronoff, Mark & Lindsay, Mark. 2014. Partial organization in languages: *La langue est un système où la plupart se tient*. In *Proceedings of the Décembrettes 8th International Conference of Morphology. Carnets de Grammaire* 22: 1–14.

Bárány, András. 2015. Hungarian object agreement with personal pronouns. In *Proceedings of the 32nd West Coast conference on formal linguistics*, Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara & Iris Chuoying Ouyang (eds), 208–217. Somerville MA: Cascadilla Proceedings Project.

Bartos, Huba. 2000. Az inflexiós jelenségek szintaktikai háttere. In *Strukturális magyar nyelvtan. 3. Morfológia*, Ferenc Kiefer (ed.), 653–762. Budapest: Akadémiai Kiadó.

Bermel, Neil & Knittl, Luděk. 2012. Morphosyntactic variation and syntactic constructions in Czech nominal declension: corpus frequency and native-speaker judgments. *Russian Linguistics* 36: 91–119. https://doi.org/10.1007/s11185-011-9083-x

Berruto, Gaetano. 1995. *Fondamenti di sociolinguistica*. Roma: Laterza.

Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana & Baayen, R. Harald. 2007. Predicting the dative alternation. In *Cognitive Foundations of Interpretation*, Gerlof Bouma, Irene Krämer & Joost Zwarts (eds), 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Cappellaro, Chiara. 2013. Overabundance in diachrony: A case study. In *The Boundaries of Pure Morphology. Diachronic and Synchronic Perspectives*, Silvio Cruschina, Martin Maiden & John Charles Smith (eds), 209–220. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199678860.003.0011

Creissels, Denis. 2006. Suffixes casuels et postpositions en hongrois. *Bulletin de la Société de Linguistique de Paris* 101: 225–272. https://doi.org/10.2143/BSL.101.1.2019827

Crysmann, Berthold & Bonami, Olivier. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52: 311–374. https://doi.org/10.1017/S0022226715000018

D. Mátai, Mária. 2003. Szófajtörténet. In *Magyar nyelvtörténet*, Jenő Kiss & Ferenc Pusztai (eds), 213–216, 398–400, 635–636, 740–741, 824. Budapest: Osiris Kiadó.

den Dikken, Marcel. 2004. Agreement and 'clause union'. In *Verb Clusters. A Study of Hungarian, German and Dutch*, Katalin É. Kiss & Henk van Riemsdijk (eds), 445–498. Amsterdam: John Benjamins. https://doi.org/10.1075/la.69.24dik

den Dikken, Marcel. 2006. When Hungarians agree (to disagree) – The fine art of 'phi' and 'art'. Ms, CUNY Graduate Center. <http://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Programs/Linguistics/Dikken/hungarian_agreement_2006.pdf> (August 2017).

É. Kiss, Katalin. 2002. *The Syntax of Hungarian*. Cambridge: CUP.
https://doi.org/10.1017/CBO9780511755088

Hunfalvy, Pál. 1856. Bennünket, benneteket. *Magyar Nyelvészet* 1: 232–236.

Kenesei, Istvan, Vago, Robert M. & Fenyvesi, Anna. 1998. *Hungarian*. London: Routledge.
https://doi.org/10.4324/9780203192238

Kiefer, Ferenc. 2000. A ragozás. In *Strukturális magyar nyelvtan, 3: Morfológia*, Ferenc Kiefer (ed.), 569–600. Budapest: Akadémiai Kiadó.

Kroch, Anthony (1994). Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistic Society. Volume 2: The Parasession on Variation in Linguistic Theory*, Katharine Beals, Jeanette Denton, Robert Knippen, Lynette Melnar, Hisami Suzuki & Erica Zeinfeld (eds), 180–201. Chicago IL: Chicago Linguistic Society.

Kypriotaki, Lyn. 1974. Acquisition of AUX. *Papers and Reports on Child Language Development* 8: 87–103.

Lindsay, Mark & Aronoff, Mark. 2013. Natural selection in self-organizing morphological systems. In *Morphology in Toulouse: Selected Proceedings of Décembrettes 7*, Fabio Montermini, Gilles Boyé & Jesse Tseng (eds), 133–153. Munich: Lincom.

Rounds, Carol. 2009. *Hungarian: An Essential Grammar*. London: Routledge. https://doi.org/10.4324/9780203465196

Spencer, Andrew. 2008. Does Hungarian have a case system? In *Case and Grammatical Relations: Studies in honor of Bernard Comrie* [Typological Studies in Language 81], Greville G. Corbett & Michael Noonan (eds), 35–56. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.81.02doe

Spencer, Andrew J. & Stump, Gregory T. 2013. Hungarian pronominal case and the dichotomy of content and form in inflectional morphology. *Natural Language and Linguistic Theory* 31: 1207–1248. https://doi.org/10.1007/s11049-013-9207-7

Steiner, Zsigmond. 1873. Bennünket, benneteket. *Magyar Nyelvőr* 2: 444–446.

Stump, Gregory. 2016. *Inflectional Paradigms. Content and Form at the Syntax-morphology Interface*. Cambridge: CUP. https://doi.org/10.1017/CBO9781316105290

Szili, Katalin. 2000. A tárgyasság a magyar nyelvben. A tárgyasság morfoszintaktikai összefüggései. *Magyar Nyelvőr* 124: 356–369.

Tálos, Endre. 1980. A magyar inkluzív személynévmásai. *Magyar Nyelv* 76: 70–72.

Thornton, Anna M. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology. In *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, Martin Maiden, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds), 358–381. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199589982.003.0017

Thornton, Anna M. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5: 183–207. https://doi.org/10.3366/word.2012.0026

Thornton, Anna M. 2016. Un capitolo di storia della terminologia grammaticale italiana: Il termine *sovrabbondante*. In *Categorie grammaticali e classi di parole. Statuto e riflessi metalinguistici*, Francesco Dedè (ed.), 289–309. Rome: Il Calamo.

Thornton, Anna M. To appear. Overabundance: A canonical typology. In *Competition in Morphology*, Franz Rainer, Francesco Gardani, Hans-Christian Luschützky & Wolfgang U. Dressler (eds). Dordrecht: Springer.

Thuilier, Juliette. 2012. *Contraintes préférentielles et ordre des mots en français*. Thèse de Doctorat, Université Paris-Diderot – Paris VII. <https://tel.archives-ouvertes.fr/tel-00781228> (10 May 2015).

Zaicz, Gábor. 2006. *Etimológiai szótár. Magyar szavak és toldalékok eredete*. Budapest: TINTA Könyvkiadó.

Corpora

HDC = <http://www.nytud.hu/hhc/pat3h.html> (June–August 2017); a new version is available at <http://clara.nytud.hu/mtsz/run.cgi/first_form>

MNSZ2 = <http://corpus.nytud.hu/mnsz/index_eng.html> (June–August 2017).

OHC = <http://omagyarkorpusz.nytud.hu/en-intro.html> (June–August 2017).

Unstable personal pronouns in Northern Logudorese

Michele Loporcaro, Serena Romagnoli and Mario Wild
University of Zurich

This paper deals with some Logudorese dialects of Northern Sardinia, whose nominal and pronominal morphology has been variously reshaped due to contact with Gallurese/Sassarese. While many of the data discussed here were addressed in the previously available literature, we draw on first-hand fieldwork data to show that further simplification in the system of personal pronouns has taken place (Sènnori) or is presently ongoing (Luras), resulting in an overall trend towards loss of marking of the gender contrast, everywhere in the plural, at times also in the singular. In particular in the case of Luras, we show that this re-arrangement is taking place through a kaleidoscope of subtly differing individual variants.

Keywords: gender/number agreement, variation (cross- and intra-dialectal), language contact, language change

1.   Introduction

As sundry studies have shown (cf. Campus 1901: 15; Bottiglioni 1920: 30 f.; Wagner 1923: 241; Jäggli 1959: 18; Sanna 1975: 106 f.; Loporcaro 2003, 2012a, 2015; Manzini & Savoia 2005), some Logudorese dialects of northern Sardinia (indicated on map 1) have restructured several aspects of morphology in ways differing from place to place, with a general trend towards simplification. This has resulted in loss of contrast in the plural between masculine and feminine agreement targets, and possibly also in loss of contrast, again in the plural, between noun inflectional classes associated with masculine vs. feminine gender. The studies that have dealt with the topic– beginning with Campus (1901)– correctly traced back the origin of these changes to the contact pressure exerted by neighbouring varieties. Due to contact with Sassarese, the dialect of Sènnori has generalized the plural ending *-ɔs*, originally only masculine, as first remarked by Campus (1901: 15):

"Il plurale dei nomi femminili in -*a* esce in -*os* come quello dei maschili in -*u*". This change, in Sennorese, invested the inflection both of nouns and of all agreement targets, including stressed personal pronouns. Analogously, the dialect of Luras, in contact with Gallurese, got rid of the gender distinction in plural agreement targets (class 1 adjectives, past participles, pronominal clitics), which all display a single gender-unmarked ending -*as*. Contrary to Sennorese, though, in the 3rd person plural pronoun, Lurese has preserved the original gender-marked forms *issɔs* and *issas* (used under particular circumstances, as we will see in § 3.2), and added to the inventory an analogical gender-neutral form *issɛs*. So far the situation described in the previous literature up to Loporcaro (2003) and Manzini & Savoia (2005: 3.462).

However, in fieldwork carried out in summer 2016, we realized that the situation, in the speech community, is today more complicated than described in the above sources, due to ongoing restructuring in the pronoun system. This is the main focus of the present paper, whose breakdown is as follows: in § 2, we briefly address for comparison the pronominal system of Sennorese, showing that the simplification it had undergone has given rise to a stable arrangement, not subject to ongoing change at present. In § 3, we move on to consider Lurese, illustrating how change in personal pronouns is progressing across the speech community, to finally (§ 4) compare the dialects of Bulzi and Perfugas, whose morphology also has undergone similar reshaping.

2.   Contact-induced change in gender agreement and noun inflection in Sennorese

Let us first consider the dialect of Sènnori. This is a Northern Logudorese variety in close contact with Sassarese: in fact, Sassarese is spoken up to the village of Sorso, which is only 2 km away from Sènnori. In other words, even though the two towns of Sènnori and Sorso are contiguous, such geographical continuity is interrupted by the linguistic border between Sassarese and Logudorese. This peculiar situation has caused contact influences in both directions and, as a result, several changes have permeated this border.

For a previous stage of Sennorese, one has to assume the common Logudorese system as schematized in (1):[1]

---

1. Dialect data are presented in a simplified IPA transcription, where word stress is notated with an acute only on non-paroxytonic words and consonant gemination is notated with the repetition of the consonant symbol. In the schemas introduced in what follows, solid lines divide paradigm cells which host contrasting forms.

(1) a. 3rd person pronoun

	SG	PL
M	iss-ɛ/-u	iss-ɔs
F	iss-a	iss-as

(Logudorese)

  b. DET and noun inflection

	SG	PL			
M	su ɣaɖɖu	sɔs kaɖɖɔs	< CABALLUM, -OS	class 2	'horse'
F	sa ɣraβa	sas kraβas	< CAPRAM, -AS	class 1	'goat'

As exemplified with 3rd person pronouns in (1a) and with definite articles in (1b), Logudorese displays parallel gender agreement (see Corbett 1991: 155), contrasting masculine and feminine in both the singular and the plural. Sennorese, on the other hand, has convergent gender marking on the definite article ((2b)) and on most other agreement targets (demonstratives, adjectives, past participles, direct object clitics), in that the plural does not contrast masculine and feminine. This is illustrated also with the generalization of one plural form *issɔs* in (2a) for stressed 3rd person pronouns[2]:

(2) a. 3rd person pronoun

	SG	PL
M = F	iss-ɛ	iss-ɔs

(Sennorese)

  b. DET and noun inflection

	SG	PL			
M	s(u) attu	sɔ vváttɔzɔ	< CATTUM, -OS	class 2	'cat'
F	s(a) akka	sɔ bbákkɔzɔ	< VACCAM, -AS	class 1	'cow'

Convergence in the plural has been achieved through generalization of the previously masculine form, as *sɔs* in (2b) (where final /s/ is not realized since it undergoes sandhi assimilation) ousted *sas*, and *issɔs* in (2a) ousted *issas*. This is further exemplified for other agreement targets in (3) (3a-b are from Manzini & Savoia 2005: 3.589):

(3) a. kuss-ɔ   bbɛll-ɔ   ffé:min-ɔ(zɔ)
        DEM.PROX-PL   beautiful-PL woman(F)-PL
        'these beautiful women'
    b. kuss-ɔ   bbɛll-ɔ   llíbbar-ɔ(zɔ)
        DEM.PROX-PL   beautiful-PL book(M)-PL
        'these beautiful books'

---

2. Data illustrating this peculiarity of the dialect of Sènnori are available in the previous literature: cf. Jäggli (1959: 18); Sanna (1975: 106); *ALI* – e.g. 1.47– in which Sènnori is pt. 710; Manzini & Savoia (2005: 3.589).

c. s-ɔ   vvakk-ɔ/sɔ vvɔ-ɛ   ll-ɔz   appɔ   énnið-ɔzɔ
    DEF-PL cow(F)-PL/ox(M)-PL   3-PL   have.PRS: 1SG sold-PL
    'the cows/oxen, I have sold them'.

Comparison with the system of Sassarese ((4)) shows that convergence in plural gender/number agreement in Sennorese is a replica of the Sassarese system. However, in the latter the plural endings merged due to regular sound change (see Guarnerio 1898: 189–191):

(4) a. 3rd person pronoun

	SG	PL
M	ɛdd-u	ɛdd-i
F	ɛdd-a	

(Sassarese)

b. DET and noun inflection

	SG	PL		
M	lu zóːrigu	li zóːrigi	class 2	'mouse'
F	la liŋga	li liŋgi	class 1	'tongue'

The examples in (2)–(3) also show that the generalization of -ɔs (to the detriment of Logudorese -as) has affected noun inflection as well, so that class 1 and 2 nouns nowadays only contrast in the singular, just like agreement targets. This too is a replica of Sassarese, where -i generalized also in class 1 nouns, which originally shared the Proto-Romance ending *-/e/ with feminine plural articles, demonstratives, participles, and class 1 adjectives.

     In our fieldwork, we have also observed that speakers have further simplified the system of stressed personal pronouns beyond the convergence found elsewhere in Sassarese (and Gallurese).[3] In fact, as shown in (2a), also the singular 3rd person pronoun is not any longer gender-inflected, and the form *issɛ* (originally masculine) is used also for feminine referents, parallel to plural. It seems as though, once initiated by language contact, the change had overshot its (Sassarese) model. Thus, in the dialect of Sènnori, stressed personal pronouns never offer speakers the option to disambiguate gender, even where this would be useful. This stands out clearly in the embarrassment of our informant, in the following exchange:

(5) Interviewer DP: Ma chi, lui o lei?
                      'but who (do you mean), him or her?'
     Informant AM: (hesitating) *ma ɣie issɛ o issɛ*? … Beh, noi diciamo sempre maschile per tutti…
                      'but who 3SG or 3SG? … well, we always say masculine to all …'

---

3. The previous literature on Sennorese, mentioned in § 2 and fn. 2, did not address personal pronouns.

Interviewer DP: Sono uguali
'they are the same'

3.   Contact-induced change in the dialect of Luras

3.1   Convergent gender marking in Lurese

Like Sennorese, also the Northern Logudorese dialect of Luras has reshaped its inflectional morphology due to contact pressure, in ways that are partly similar but in part differ interestingly from Sennorese. For Lurese too, the starting point, to be assumed for an earlier stage, is the common Logudorese system in (1), which in this case changed due to contact with Gallurese, whose system (identical to Sassarese's in the relevant respects) is schematized in (6)[4]:

(6) a.   3rd person pronoun

	SG	PL
M	idd-u	idd-i
F	idd-a	

(Gallurese)

b.   DET and noun inflection

	SG	PL			
M	lu jattu	li jatti	< CATTUM,-I/-OS	class 2	'cat'
F	la akka	li akki	< VACCAM,-AE/-AS	class 1	'cow'

Gallurese too displays a general convergence with systematic neutralization of the masculine/feminine contrast in the plural, which has affected nouns, pronouns and agreement targets of all word classes. Again, as in Sassarese, this has a phonetic reason, i.e. the raising of final -E > -*i*. Consider now Lurese:

(7) a.   gender/number agreement

	SG	PL
M	-u	-as
F	-a	

(Lurese)

b.   DET and noun inflection

	SG	PL			
M	su ɣaɖɖu	sas kaɖɖɔs	< CABALLUM, -OS	class 2	'horse'
F	sa véːmina	sas féːminas	< FEMINAM, -AS	class 1	'woman'

Much like Sennorese, this Logudorese dialect– whose speakers are all bidialectal with Gallurese– has replicated the same agreement convergence in spite of the

---

4. Data from Guarnerio (1898: 189, 192). In (7b), the lack of the divide in the plural is motivated by the neutralized form of the articles (even if the nouns still show distinct inflections).

phonetic reason being missing. Symmetrically with respect to Sennorese, the plural ending previously marking the feminine plural has now been extended to serve as a mere plural ending for all agreement targets. This is further exemplified in (8):

(8) a. (sas káɖɖɔzɔ) laz/*lɔz appɔ ɣɔmpará:ðaza/*ɣɔmpará:ðɔzɔ
'(the horses) I have bought them'

b. (saⁱlʲ vákkaza) laz/*lɔz appɔ ɣɔmpará:ðaza/*ɣɔmpará:ðɔzɔ
'(the cows) I have bought them'.

The examples in (7)–(8) also show that, contrary to Sennorese, noun inflection remained unaffected by this change, as class 1 and class 2 nouns preserve distinct inflections not only in the singular, but also in the plural.

3.2 Change in the personal pronoun system: From common Logudorese to conservative Lurese

While the Lurese data in (7)–(8) have been addressed in the literature mentioned in § 1 (mostly exemplifying convergence with the definite article), Manzini & Savoia (2005) provide a richer set of examples. However, in the context of describing convergent agreement, Manzini & Savoia (2005: 3.648) do not discuss stressed personal pronouns, and the examples adduced for *isse* throughout their three volumes (1.729, 2.40) are always glossed as 'lui' (3M.SG) (*s (iss) isse nɔ bbéni:ði el méddzuzu* "se (lui) lui non viene è meglio" 1.729, 'if he does not come, it's better', where the first *lui* literally renders the optional final part of the hypothetic conjunction (stemming from a pronominal form, etymologically), while the second one is a 3M.SG pronoun). Similarly, the examples of the 3PL pronoun (*íssεzε* in Luras) are glossed with Italian ambigeneric *loro* (see 1.516, 2.344, 3.462, 3.482):

(9) čama:ðε   a íssεzε
call:IMPER.2PL to 3PL
'call them'.
        (Luras; glossed "chiamate loro" in Manzini & Savoia 2005: 3.462, 3.482)

Thus, on the whole, one could get the impression that the stressed personal pronoun system here is identical to that of Standard Italian. However, the Lurese dictionary by Depperu (2006: 389) records not only ambigeneric "**Isses** *pron. pers.m.&f.pl. a)* Essi, loro; *b)* Esse, loro", but also "**Issas** *pron.pers.f.pl* Loro, esse". Furthermore, the author of the dictionary was among the Lurese speakers I interviewed in the 2003 fieldwork session: there, I ascertained that his competence, much like that of the other elderly speakers I then interviewed, actually reflected the system schematized in (10a), henceforth referred to as "conservative Lurese" for reasons to become clear shortly:

(10)   3SG/PL stressed pronouns in Luras (Loporcaro 2015: 116)

a.

	SG	PL	
M = F	issɛ	issɛs	(gender-unmarked)
M	issu	issɔs	(masculine)
F	issa	issas	(feminine)

b. Latin etyma

< | IPSE | --- |
    |-------|-------|
    | IPSUM | IPSOS |
    | IPSAM | IPSAS |

First of all, one finds the forms *issɛ* (singular) and *issɛs* (plural), the former occurring elsewhere in Logudoro as 3M.SG, while the plural one was shaped in this dialect by analogy on the singular, as suggested by the lack of an etymological source, highlighted in (10b). These two (*issɛ* and *issɛs*) are the forms used in ordinary conversation to refer to 3rd person human referents.[5] Alongside these gender-unmarked forms, however, the marked ones (*issu, -a, -ɔs, -as*) survive into this system, and can be used, if needed, for purposes of disambiguation:

(11)   a.   A.   ɛⁱlʲ veniṇdɛ issɛ B: ma ɣiɛ   iss-u   ɔ iss-a?
                  is coming 3SG   but who 3-M.SG or 3-F.SG
                  A: 's/he is coming' B: 'but who (do you mean)? Him or her?'
       b.   kiɛ bb' ɛstɛ / kiɛ aⁱlʲ   viðu, iss-ɔz   ɔ íss-aza?
            who LOC is   who have: 2SG seen, 3-M.PL or 3-F.PL
            'who is there/whom have you seen, them.M or them.F?'

(Loporcaro 2006: 134)

Whenever reference is ambiguous in the context, the interlocutor has the option, in Luras, of asking a clarification question using gender marked forms.

3.3   Instability and ongoing change in personal pronouns in the dialect of Luras

While the convergence in gender agreement on all remaining agreement targets, seen in § 3.1, is represented in the dialect of all Lurese speakers, the informants we interviewed in 2016 differ in their use of 3rd person stressed pronouns. In fact, our fieldwork has revealed a tendency towards simplification of the traditional system in (10a)– repeated in (12a) with the abbreviations indicating the informants recorded in 2003 and in 2016 (plus their birth year). At present, this

---

5. Another instance of contact-induced change in this area of grammar is the borrowing of a 3rd person pronoun form *ɛḍḍɛ*, which cannot be autochthonous in Logudorese since it contains the pronoun stem ILL(E) instead of IPS(E) (with a stressed vowel corresponding to Sassarese, rather than Gallurese: see (6) and (4)). Several Lurese speakers use this borrowed form, restricted to the phrase *da βɛr ɛḍḍɛ* 'on his/her own', as an alternative to autochthonous *da βɛr issɛ*. While the stem is borrowed, the ending is reshaped in accordance with the Lurese personal pronoun system (10a).

pattern represents the conservative option in the arrangement of the stressed pronoun system.

(12)   3rd person stressed pronouns in the dialect of Luras: cross-individual variation

a. conservative (stage 1)   b. intermediate A (stage 2)   c. innovative (stage 3)

	SG	PL
M = F	issɛ	issɛs
M	issu	issɔs
F	issa	issas

SA (1918), PD (1943), FS (1938),DS (1957)

	SG	PL
M = F	issɛ	issɛs
M	issu	
F	issa	issas

AL (1959), PL (1942), FD (1943), MGA (1947), MM (1960)

	SG	PL
M = F	issɛ	issɛs
M	issu	
F	issa	

GS (1986), LDS (1991)

Alongside the conservative system in (12a), it is possible to trace, in today's Lurese, the two further systems observed in the speakers indicated in (12b-c). Taken together, (12a-c) are easily interpreted as three subsequent stages of an ongoing change. At the endpoint of change– stage 3 in (12c)– one finds the youngest speakers interviewed in 2016. The innovative pronoun system shows that Lurese is heading towards full neutralization in the plural. This simplification, again, replicates the Gallurese system, which in turn coincides in this with Italian. In the singular, on the other hand, the gender-marked forms are preserved, and remain available for disambiguation. In between, one finds the intermediate system in (12b) (stage 2), displayed by most informants today. Compared with the conservative system ((12a)), these informants have lost the M.PL form *issɔs and hence have a binary contrast in the plural. As a result, they select issɛs for both gender-unmarked use and for the masculine (contrasting with issas) when needed for disambiguation.

Yet another option, which we observed in just one speaker ((13a)), shows retention of marked M.PL issɔs, a different "intermediate" system with respect to which the further system in (13b)– also observed in just one informant– is likely to be a further development:

(13)   3rd person stressed pronouns in the dialect of Luras: cross-individual variation 2

a. intermediate B

	SG	PL
M = F	issɛ	issɛs
F	issa	
M	issu	issɔs

GC (1939)

b.

	SG	PL
M = F	issɛ	issɛs
M	issu	issɔs

GL (1939)

During the recording session, we observed not only the disappearance of the form
*issas (13a), but also a certain instability in the singular. If our understanding is
correct, this would result in (13b) in a parallel binary system in which issɛ is em-
ployed both as gender-unmarked and as F.SG contrasting with masculine issu in
disambiguation contexts.

Let us go back to the intermediate option schematized in (12b), as that is the
one displayed by the majority of our Lurese informants. The fact that these infor-
mants have lost *issɔs (rather than issas, the minority option seen in (13)), can be
explained by invoking the changes that have affected the remaining agreement
targets, in which, as said in § 3.1, one witnesses the extension of the (originally
F.PL) ending -as. Nevertheless, the final outcome here is not a generalization of
the form issas, i.e. one does not observe a blind replacement of the ending -ɔs with
-as: in fact, our informants never use *issas for masculine referents. Rather, what is
observed here is the demise of the ending -ɔs, arguably put under pressure because
it had grown less frequent, being retained only in noun inflection but expunged
from all other word classes in Lurese. Grammaticality judgements by our infor-
mants in (12b) confirm their perception of the ending -ɔs as non-autochthonous,
as seen in the metalinguistic commentary by one of our informants (MM 1960):

> allora, *issɔzɔ forse lo dicono poi nell'oschirese …, quelle zone lì, perché molte
> cose che nel resto del Logudoro, anche a Olbia, si dicono al maschile, noi le di-
> ciamo al femminile
> [well, issɔzɔ, perhaps they say it in the dialect of Oschiri …, (in) that area, because
> many things that in the rest of Logudoro, and in Olbia too, are said in the mascu-
> line, we say in the feminine]

This judgement is indicative of subjective perception, and interestingly shows a
lack of awareness of the fact that speakers representative for the conservative sys-
tem (12a), with whom MM 1960 happens to interact, are still found within the
same speech community.

The difference between what is happening in stressed pronouns, in (12b-c),
and the way convergence is achieved on other agreement targets ((7)–(8), § 3.1)
can be explained with reference to the particular status of stressed pronouns. In
fact, these are not just mere agreement targets: they can be either agreement tar-
gets or agreement controllers when used in their capacities as anaphoric or deictic
pronouns respectively. In addition, gender on 3rd person pronouns is referentially
motivated, given that, not unlike Italian lei/lui/loro, Lurese (and, more generally,
Sardinian) stressed 3rd person pronouns are used only for human referents. By
contrast, gender agreement with the other targets considered in § 3.1, though
found with a semantic core (nouns denoting male/female humans and 'higher'
animals), is overwhelmingly determined by lexical idiosyncrasy, as is generally

the case in the Romance languages. In addition, the fact that in (12b) only the F.PL form *issas* can be used for disambiguation, thus coexisting with unmarked *issɛs*, corresponds to the general situation to be found in the Romance languages with binary gender systems, where the masculine can be used for gender-generic reference (this also goes for Italian *essi*, the masculine 3PL form, alternative to most common *loro*).

4.   Personal pronouns in the nearby dialects of Bulzi and Pèrfugas

The dialects of Bulzi and Pèrfugas display a situation partly similar to that of Lurese, as pointed out by Maxia (2010: 143) with regard to the definite article:

> Alcuni centri situati lungo la linea di contatto tra il sardo logudorese e le varietà sardo-corse presentano l'articolo unificato *sos* 'i, gli, le' (Sennori, Olbia) oppure *sas* 'i, gli, le' (Bulzi, Perfugas, Luras).
> [Some centres located along the line of contact between Logudorese Sardinian and the Sardinian-Corsican varieties display the unified article *sos* 'the.M.PL=F.PL' (Sennori, Olbia) or *sas* the.M.PL=F.PL' (Bulzi, Perfugas, Luras).]

In the dialect of Bulzi (*ALI*, pt. 713), in noun inflection the contrast between plural *-ɔs* (in class 2 nouns, overwhelmingly masculine) and plural *-as* (in class 1 nouns, prevalently feminine) is retained, whereas one observes convergence in *-as*, as in Lurese, not only in the definite article but also in demonstratives and pronominal clitics, class 1 adjectives and past participles:

(14) a.   s-az   amiːɣ-as   sáld-aza
           DEF-PL   friend(F)-PL Sardinian-PL
           'the Sardinian (female) friends'
      b.   s-az   amiːɣ-ɔs   sáld-aza
           DEF-PL   friend(M)-PL   Sardinian-PL
           'the Sardinian friends'.

As far as stressed personal pronouns are concerned, among the few informants that we interviewed in 2016 we observed two transitional systems identical to the two reported above for Lurese in (12b) and (13b). Here too, the singular shows a certain instability, while the general trend is towards simplification, with generalization of *issɛs* in the plural:

(15)   3rd person stressed pronouns in the dialect of Bulzi

a.

	SG	PL
M = F	issɛ	isses
M	issu	
F	issa	issas

RS (1942)

b.

	SG	PL
M = F	issɛ	issɛs
M	issu	issɔs

LC (1956)

In Perfugas too, the system seems to be undergoing change. Here, adjective inflec-tion is parallel to noun inflection, with the contrast between M.PL *-ɔs* and F.PL *-as* preserved. The definite article, instead, is now shifting towards the extension of the originally F.PL form *sas* also to M.PL[6]:

(16)   a.   s-az   amiːɣ-as   sáld-aza
            DEF-PL   friend(F)-PL Sardinian-F.PL
            'the Sardinian (female) friends'
      b.   s-az   amiːɣ-ɔs   sáld-ɔzɔ
            DEF-PL friend(M)-PL Sardinian-M.PL
            'the Sardinian (male) friends'.

This convergence is observed variably also in personal pronouns, as shown by the two different systems in (17a-b), the former parallel, the latter convergent:

(17)   3rd person stressed pronouns in the dialect of Perfugas: cross-individual variation

a.

	SG	PL
M	issu	issɔs
F	issa	issas

SP (1946)

b.

	SG	PL
	issu	issɛs
	issa	

ES (1934)

In fact, the speaker born in 1934 (17b) only accepts the form *isses* for the plu-ral, whereas in the singular he has the two gender-marked forms *issa* and *issu*. Notwithstanding the lack, in this speaker, of the singular form **issɛ*, a preceding stage with a common Logudorese type of system (as in the schema in (1a)) is sug-gested by the occurrence of plural *isses*: in fact, the creation of the latter– as seen above in (10) for Lurese– must have happened by analogy on singular **issɛ*. Perhaps also here, as in Luras, *issɛ* may have been refunctionalized in an intermediate stage as a gender-unmarked pronoun, so that plural *isses* was created subsequently on its model, before singular **issɛ* went out of use. If this reconstruction is on the right track, this speaker would attest to a subsequent stage in which the system has been

---

6. The originally M.PL form *ɔs* is deemed as only marginally grammatical by our informants, and its distribution does not seem to respect any phonological criterion.

simplified both in the plural (with retention of only of the gender-unmarked form) and the singular (with retention of both gender-marked forms). As an alternative, one could interpret *issɛs* as a loanword from a neighbouring dialect.[7]

A somewhat younger speaker we have interviewed, whose system is schematized in (17a), shares this schema as far as the singular is concerned, while the situation in the plural contrasts diametrically: this informant presents the same schema seen for common Logudorese in (1), with the *issɔs* vs. *issas* contrast, (while *issɛs* is deemed ungrammatical). If this situation is, in Perfugas, a secondary development– as the age of the informant would seem to suggest – we would be facing a contact-induced restoration of the common Logudorese system, still alive and well in the Logudorese dialects spoken further south.

At any event, the empirical basis is still limited, and more reliable conclusions on this point will be possibly reached thanks to the dedicated fieldwork on the dialects of Bulzi and Perfugas foreseen in the framework of the ongoing study by Wild (2017).

5.   Conclusion

The data discussed lend themselves to some remarks about method, on several fronts: linguistic-typological generalizations, structural-analytical aspects and sociolinguistic considerations. On the first front, this study clearly confirms that, as soon as the dialects of Italy are considered in-depth, typologically interesting facts emerge, as is the case for the pronominal system of Lurese, not inventoried for any of the world's languages before our fieldwork: as remarked in Loporcaro (2015: 119), the Lurese system would require adding a new colour in the WALS map on gender marking on the personal pronoun (Siewierska 2005).

From a structural-analytical point of view, the overall picture of contact-induced changes in the morphology and in the morphosyntax of the varieties under examination shows that it is useful to keep the two aspects apart (in particular for the noun), distinguishing between gender (morphosyntax) and inflectional class (pure morphology). The changes in the two areas may proceed jointly, as in Sènnori, where feminine nouns have changed the plural ending on a par with all agreement targets, including stressed personal pronouns. On the other hand, change may also concern only the morphosyntax of agreement, sparing noun morphology, as happened in Luras and Bulzi. To the litmus test of these changes,

---

7. The further data gathered in Wild (2017) seem to favour this interpretation of ES's system. In this particular case the neighbouring dialect inducing the change in this single speaker's grammar can be identified in Bulzese, to which he is exposed in the person of his wife.

stressed personal pronouns reveal their particular status (as both agreement controllers and targets) in that they may follow the destiny of the other word classes with similar inflections, as is the case for the plural in Sennorese, or else go autonomous paths, as in Bulzi, Perfugas and Luras.

Such autonomous paths– and this is our concluding remark, sociolinguistic in nature– may show a fine-grained differentiation across speakers in the speech community, with outcomes diffracted to an extent unknown to standard languages that are subject to normative pressure. Whether this picture necessarily ensues from the sociolinguistic nature of a dialect (as opposed to a standardized language) or whether this represents instead– for the dialects of today's Italy– an instance of "fringing" of the kind familiar from studies on terminal language stages preceding language death in a context of generalized language shift (see e.g. Dorian 1981, among a vast number of studies on language death), is a question that, based on the data discussed in the present article, we can here raise, but not answer.[8]

Acknowledgments

---

8. That language shift is ongoing, and at a speedy pace, throughout most of Sardinia, is an incontestable fact (see the sociolinguistic study in Oppo 2007 and the discussion in Loporcaro 2012b). However, in our case, given the sociolinguistic profile of the small centres investigated here (tightly-knit communities, which generally tend to favour the maintenance of linguistic complexity, as argued by Trudgill 2004, 2009), informants above 50 still use Sardinian regularly in everyday life, which seems to disfavour an explanation invoking language attrition in the context of impending language death.

References

ALI = Bartoli, Matteo Giulio, Vidossi, Giuseppe, Terracini, Benvenuto A., Bonfante, Giuliano, Grassi, Corrado, Genre, Arturo & Massobrio, Lorenzo (eds). 1995. *Atlante linguistico italiano*. Rome: Istituto Poligrafico e Zecca dello Stato/Libreria dello Stato.

Bottiglioni, Gino. 1920. Saggio di fonetica sarda. *Studj Romanzi* 15: 5–114.

Campus, Giovanni. 1901. *Fonetica del dialetto logudorese*. Torino: Bona.

Corbett, Greville G.. 1991. *Gender*. Cambridge: Cambridge University Press.

Depperu, Pietro. 2006. *Vocabbolariu lurisincu. Dizionario logudorese della parlata di Luras*. Sassari: Libreria Koinè.

Dorian, Nancy. 1981. *Language Death. The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia PA: University of Pennsylvania Press. https://doi.org/10.9783/9781512815580

Guarnerio, Pier Enea. 1898. I dialetti odierni di Sassari, della Gallura e della Corsica (continuazione). *Archivio Glottologico Italiano* 14: 131–200, 385–422.

Jäggli, Peter. 1959. *Die Mundart von Sénnori (Provinz Sassari, Sardinien). Ein Beitrag zur Kenntnis der nordlogudoresischen Mundarten*. Zurich: Juris-Verlag.

Loporcaro, Michele. 2003. Dialettologia, linguistica storica e riflessione grammaticale nella romanistica del Duemila. Con esempi dal sardo. In *Actas del XII Congreso Internacional de Lingüística y Filología Románica*, Salamanca, 24–30 septiembre 2001, Fernando Sánchez Miret (ed.), 83–112. Tübingen: Max Niemeyer.

Loporcaro, Michele. 2006. Contatto e mutamento linguistico in Sardegna settentrionale: Il caso di Luras. *Revue de Linguistique Romane* 70: 321–349

Loporcaro, Michele. 2012a. Contact-induced change in personal pronouns: Some Romance examples. In *Morphologies in Contact* [Studia Typologica, STUF Beiheft 10], Martine Vanhove, Thomas Stolz, Aina Urdze & Hitomi Otsuka et al. (eds), 207–225 Berlin: De Gruyter. https://doi.org/10.1515/9783110399967.

Loporcaro, Michele. 2012b. Non sappiamo come scriverlo, perciò non lo parliamo: Mille e una scusa per un suicidio linguistico [Non ischimos comente l'iscrier e tando sensamos de lu faeddare: Milli e una corrighetta pro bocchire una limba]. *Rhesis. International Journal of Linguistics, Philology, and Literature. Quaderni del Dipartimento di Linguistica e Stilistica dell'Università di Cagliari* 3: 36–58 <http://www.diplist.it/rhesis/index.php?phpsessid=8886c348ec8531816aeb78bcc30b6316&sez=archiveLP> (6 February 2018).

Loporcaro, Michele. 2015. The impact of morphology on change in agreement systems. In *Agreement from a Diachronic Perspective*, Jürg Fleischer, Elisabeth Rieken & Paul Widmer et al. (eds), 104–126. Berlin: De Gruyter. https://doi.org/10.1515/9783110399967.

Manzini, M. Rita & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni dell'Orso.

Maxia, Mauro. 2010. *Studi sardo-corsi. Dialettologia e storia della lingua tra le due isole*. Olbia: Taphros.

Oppo, Anna (ed.). 2007. Giovanni Lupinu, Alessandro Mongili, Anna Oppo, Riccardo Spiga, Sabrina Perra & Matteo Valdes, *Le lingue dei sardi. Una ricerca sociolinguistica*. Cagliari-Sassari: Regione autonoma della Sardegna, Università di Cagliari, Università di Sassari.

Sanna, Antonio. 1975. *Il dialetto di Sassari*. Cagliari: Edizioni 3T.

Siewierska, Anna. 2005. Gender distinctions in independent personal pronouns. *WALS* 44: 182–185.

Trudgill, Peter. 2004. The impact of language contact and social structure on linguistic structure. In *Dialectology Meets Typology. Dialect Grammar from a Cross-linguistic Perspective*, Bernd Kortmann (ed.), 435–451. Berlin: Mouton de Gruyter.

Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In *Language Complexity as an Evolving Variable*, Geoffrey Sampson, David Gil & Peter Trudgill (eds), 98–109. Oxford: OUP.

Wagner, Max Leopold. 1923. Zur Stellung des Galluresisch-Sassaresischen (Aus Anlaß von Bottiglionis 'Saggio di fonetica sarda'). *Archiv für das Studium der neueren Sprachen und Literaturen* 145: 239–249; 146: 98–112, 223–228.

WALS = Haspelmath, Martin, Dryer, Matthew S., Gil, David & Comrie, Bernard (eds). 2005. *The World Atlas of Language Structures*. Oxford: OUP.

Wild, Mario. 2017. Note di morfologia e morfosintassi su quattro dialetti logudoresi settentrionali. I casi di Sennori (SS), Perfugas (SS), Bulzi (SS) e Nulvi (SS). MA thesis, University of Zurich.

Object clitics for subject clitics in Francoprovençal and Piedmontese

Ian Roberts

Downing College, University of Cambridge

This article illustrates and analyses the intricate phenomenon of OCL-for-SCL, found in certain varieties of Franco-Provençal Valdôtain and Piedmontese. This phenomenon, which at first sight appears highly unusual, reflects operations of morphophonological realisation of the kind developed in the context of the theory of Distributed Morphology such as fission and fusion, amply attested elsewhere in the world's languages. In a further section, the contexts for enclisis to past participles are discussed. Following Roberts (2016), this leads to the proposal of the implicational scale 'complement to restructuring verb > complement to auxiliary > complement to causative'. This can be described this in terms of degrees of 'transparency' which in turn may translate into structural 'size'.

Keywords: Clitics, enclisis, fission, fusion, distributed morphology

1.   Introduction

Certain varieties of Valdôtain, Piedmontese, and Friulian show a phenomenon which Roberts (1993: 329f.) calls 'OCL-for–SCL':

(1) a. Gnunc l'   a   viu -me   (Valdôtain of Ayas)
         noone 3SG.SCL= has seen = 1SG.ACC
     b. Gnunc m'   a   viu.   (Valdôtain of Ayas)
         noone 1SG.OCL= has seen
         'No-one has seen me.'

The generalization seems to be that an auxiliary always requires a proclitic; if this is a subject clitic, the object clitic is enclitic to the past participle, but many varieties allow or require the object clitic to be proclitic on the auxiliary, in which case the subject clitic does not appear.

The phenomenon in (1) raises three questions. First, what is the nature of the requirement for a pre-auxiliary clitic? Second, how do the object clitics (OCL henceforth) and the subject clitics (SCL henceforth) interact? Third, what allows enclisis of the OCL to the participle in examples like (1a)?

2.   More data

Parry (1994: 4f.) observes that in the Piedmontese of Cairo Montenotte (Cairese) the OCL appears in both positions:

(2) a. a   t   l   ø   di t   le   (Cairese)
         SCL 2SG.DAT= 3SGM.OCL I.have said = 2SG.DAT = 3SGM.ACC
         'I have said it to you.'

     b. a   ɹ   ø   vardɐ   ɹa   (Cairese)
         SCL 3SGF.ACC= I.HAVE watched = 3SGF.ACC
         'I have looked at her.'

Here we see that the auxiliary clitic follows the vocalic subject clitic (Valdôtain doesn't have vocalic clitics). Further confirmation comes from the Piedmontese of Fara Novarese (Manzini & Savoia 2005, II: 582):

(3) i   m   ɜŋ   tʃama -mi
     3PL.SCL= 1SG.ACC= they.HAVE called = 1SG.ACC
     'They have called me.'   (Piedmontese of Fara Novarese)

Also in Fara Novarese, the 3SG SCL changes form when a proclitic object clitic is present (from what Poletto & Tortora 2016 refer to as a person clitic to what is either a deictic or an invariant clitic as defined there). Manzini and Savoia refer to these as 'l-forms' and 'undifferentiated clitics' respectively:

(4) a. (a)l   ɛ   tʃama -mi   (Piedmontese of Fara Novarese)
         3SG.SCL= he.IS called = 1SG.ACC

     b. a   m   a   tʃama -mi
         3SG.SCL= 1SG.ACC= he.HAS called = 1SG.ACC
         'He has called me.'   (Piedmontese of Fara Novarese)

This can be seen as an aspect of OCL-for-SCL in that the l-form of the SCL competes with an l-form OCL. In those cases, as Manzini and Savoia (2005, II: 582) point out, the SCL is realized just as the vocalic form. It seems clear, then, that there are at least two pre-auxiliary positions for subject clitics: a higher vocalic one and a lower l-form one. Following Poletto (2000), I take the vocalic clitics to be in the left periphery of the clause, more technically the C-system. The apparent

'competition' in the lower position involves the agreement system, more exactly Person features.

In Piedmontese of Moncalvo we find simultaneous proclisis and enclisis of the OCL but without an overt SCL where the object clitic is first/second person and with an optional OCL where the object clitic is third person:

(5) a. m   aŋ   tʃama -mi
         1SG.ACC= they.HAVE called = 1SG.ACC
         'They have called me.'   (Piedmontese of Moncalvo)

     b. (al)   l   aŋ   tʃama -lu
         3SG.SCL= 3SG.ACC= they.HAVE called = 3SG.ACC
         'They have called him.'   (Piedmontese of Moncalvo)

See also Manzini & Savoia (2005, II: 440–49) for a further range of Piedmontese varieties in which the SCL consistently acts either as a person or as a number clitic (or, seemingly, as a combination of the two) in Poletto and Tortora's terms and there is no OCL proclisis.

OCL enclisis to the past participle in compound tenses appears then to be quite common in Piedmontese, with variation in whether the OCL is simultaneously proclitic to the auxiliary, with a possible effect on the appearance or form of the SCL and variation in the person of the OCL which undergoes these options (as well as sensitivity to auxiliary selection, which I leave aside here).

The interactions of OCL and SCL found here and elsewhere pose a challenge to most accounts of clitics and agreement, but seem at the very least to show coherent patterns. In particular, the fact that there is nearly always, even where clitics are enclitic to past participles, an interaction with the position proclitic to the auxiliary (in the form of copying or OCL-for-SCL) suggests that there is a clitic dependency concerning both the enclitic and the proclitic positions.

3. Analysis

3.1 Cliticisation and incorporation

In this section we will sketch out a formal analysis of OCL-for-SCL, based on the general proposals for clitics in Roberts (2010). The basic idea is a modification of the definition of 'minimal category' in the theory of bare phrase structure put forward in Chomsky (1995):

(6) A category is minimal iff it dominates no category β distinct from itself.

Consider the abstract structure in (7) in the light of the definition in (6):

(7) $[_{Y2}$ X Y_1]

Here Y_1 and Y_2 are two occurrences ('segments' in the technical terminology) of a head category, to which X has adjoined. The numerical subscripts are intended merely to distinguish the two occurrences, and have no formal status. By the definition in (6), Y_2 can be minimal, but only if X is minimal and is non-distinct from Y. So, if clitics are minimal categories (as originally proposed by Muysken 1982), X could be a clitic, as long as its category is not distinct from that of Y.

A category is just a bundle of features. So X's features cannot be distinct from Y's in (7). Therefore X's features must be a subset of Y's features. According to Cardinaletti and Starke (1999) and Déchaîne and Wiltschko (2002), Romance pronominal clitics are φPs, phrases headed by a bundle of 'φ-features' (where 'φ' is a cover term referring to Person, Number, Gender and Case). The Romance clitic pronouns differ from the tonic, 'strong' complement pronouns of a language like English in being φs, rather than Ds.

Now let us look at a more concrete case of the abstract structure in (7):

(8) $[_v$ *le* $[_v$ $[_V$ *voit*] *v*]]]

Here *v* is the abstract "light verb" associated with all transitive lexical verbs (Vs). Little *v* has two formal features, a categorial verbal feature V and a set of φ-features. The clitic pronoun *le* is composed of φ-features, as we just proposed for clitic pronouns in general. Treating formal features as Attribute-Value pairs, its features are [Pers: *3*, Num: *Sg*, Gen: *m*]. These are valued features, in the simple sense that each Attribute (Person, etc.) has a value (*3*, etc.). As such they can be interpreted, both by the semantic component (giving rise to a particular kind of pronominal reference) and by the morphophonological component (giving rise to a particular phonological shape). The *v* category has unvalued, and therefore uninterpretable counterparts of these features: [Pers:__], [Num:__], [Gen:__]. So the active, transitive *v* contains φ-features, in fact, unvalued versions of the very φ-features that make up the clitic, so the clitic is not distinct from *v*. More precisely, the clitic's φ-features form a proper subset of *v*'s features, in that *v* has features which the clitic lacks, while the clitic has no features which *v* lacks. Given the definition in (6), the clitic can adjoin to *v* and form a derived minimal head. The verb (V) does the same, as *v* is clearly categorially non-distinct from it.

This approach has two main consequences. First, 'incorporation' of one head into another, as seen abstractly in (7) and more concretely in (8), can take place only where the features of the incorporee are properly included in those of the incorporation host. The second consequence follows from the formal relation of agreement, known as Agree since Chomsky (2000). This is a relation between a Probe, a category with unvalued features which needs to find values for them in

order for them to be interpretable by the semantics and the morphophonology, and Goal, a category in the relevant structural relation to the Probe (whose precise technical formulation is not necessary here) which bears valued counterparts to the Probe's features. Where a Probe and a Goal Agree, the Goal's features are copied into the blank value-slots of the Probe's features. Now, in the situation where the Goal's features are a subset of those of the Probe, copying the values of the features of the Goal exhausts the content of the Goal. Such a Goal is defined as a defective Goal. Given this, the operation is not distinguishable from the copying involved in movement. Agreement and Movement are formally indistinguishable just where the Goal is defective, with the consequence for the morphophonology that the features are realised as part of the head constituting the Probe. The effect of cliticisation, a subspecies of head-to-head incorporation on this view, arises from the nature of defective Goals, given the technical assumptions made above.

Let us look again at a transitive vP of the kind in (8):

(9) $[_{v\text{P}} \ [_v \ \varphi \ [_v \ \text{V} \ [_v \ \text{V}]]] \ [_{v\text{P}} \ (\text{V}) \quad (\varphi)]]$

Here, the "copy" of the object clitic "(φ)" is deleted by the usual operation which deletes copies of moved material, as in a simple case of wh-movement in English such as *who did you see (who)?* So we get the effect of movement from Agree. In standard cases of Agree (such as subject-verb agreement in English), this generally does not happen, since normally not all the features of Goal are shared with the Probe. But, precisely in this case, this ***is*** what happens. (The same applies to V here too).

Considering clitic and fully nominal direct objects, then, logically we have four options:

(10)   a.   Jean la voit.
       b.   Jean voit Marie.
       c.   * Jean voit la.
       d.   * Jean Marie-voit.

In (10a), we have cliticisation of the pronoun *la* to v, as just described. In (10b), the fully nominal direct object does not cliticise, as it has features not shared with v (e.g. D). It nonetheless abstractly Agrees with v, something not generally morphologically realised in Romance languages, except in some cases of past-participle agreement, but found in many of the world's languages. (10c) is ungrammatical because the wrong copy of *la* is phonologically realised (in general, only leftmost, structurally 'higher' copies are overtly realised). Finally, (10d) is ungrammatical because full nominals (DPs) cannot cliticise, owing to the fact that they have a D-feature which v lacks, and hence do not have a subset of v's features.

In order to understand the OCL-for-SCL phenomenon, we need to look at the structure of a periphrastic tense. This is shown in (11):

(11)   … [$_{AuxP}$ [$_{Aux}$ Aux] [$_{PartP}$ Part [$_{vP}$ EA v VP]]]

At the vP level, everything is as described above: owing to its defective nature, a complement clitic first-merged in the direct-object position in VP incorporates with v. V also incorporates to this position, and the structure in (9) results, where the clitic values v's φ-features. Next, v incorporates with Part; v is a defective goal in relation to Part as its formal features are a subset of those of Part (v's φ-features are not relevant, as these have now been valued by the clitic). So v incorporates with Part, giving rise to the following structure:

(12)   [$_{Part}$ [$_{v}$ φ v] Part]

Aux, however, has unvalued φ-features which can probe the clitic's features. The clitic now incorporates with Aux and values its features. The result of this incorporation is that the clitic and the auxiliary act as a unit for subsequent movement operations. The resulting structure is shown in (13) (for the Italian temporal periphrasis *L'ha vista* 'S/he has seen her'):

(13)   [$_{AuxP}$ [$_{Aux}$ [$_{φ}$ *la*] [$_{Aux}$ *ha*]] [$_{PartP}$ [$_{Part}$ [$_{v}$ [$_{V}$ *vis*]] [$_{Part}$ – *ta*]]]

Let us now consider the interaction of Agree/incorporation of the clitic with the subject and its φ-features. The subject is base-generated (or "first-merged" in the terminology of bare phrase structure) in the specifier of position of vP. It values the φ-features of finite Tense (T) (this captures the fact that subject agreement is sensitive to the finiteness of the clause in many languages, including most Romance languages and English). The subject raises from the specifier position of vP (Spec,vP) to the specifier position of TP (SpecTP). The subject does not transit through Spec,PartP, so at the point in the derivation where Aux is merged, the subject is still in Spec$v$P, and as such is further away from Aux than the clitic. Hence the clitic can and must Agree and incorporate with Aux.

Aux has a Tense feature, which is responsible for attracting auxiliaries to T. The subject-agreement inflection that appears on the auxiliary, which is often morphologically fused with tense-marking, must then be associated with the T feature. So Aux bears two distinct sets of features: 'object' φ-features (actually the bundle [Pers:--, Num: --, Gen: --], as we saw above) and T-features (consisting of the Tense features themselves, however exactly these are to be specified). The T-features are valued by T, and T has φ-features which are valued by the subject in the usual way. These are the φ-features which are realised as agreement on the verb or auxiliary.

Concerning subject clitics (SCLs), which often double the agreement marking on the finite verb or auxiliary (a phenomenon known as *raddoppiamento dell'accordo* in the traditional Italian dialectological literature), Roberts (2014) proposes SCLs result from a morphophonological process of fission (see Halle & Marantz 1993): 'splitting off' of (a subset of) features of a head to form a separate morpheme. Northern Italian dialects vary significantly as to which subsets, as is well-known.

The structure of a derived tensed auxiliary before fission creates the SCL is as in (14):

(14) $[_T [_{Aux} OCL [_{Aux} AUX]] T + \varphi]]$

The SCL results from fission of T's φ-features. The OCL + AUX combination is derived as described above.

Now we are in a position to start to understand OCL for SCL: this phenomenon must derive from OCL blocking fission somehow. Roberts (2014) suggested that copying of the subject's φ-features could precede or follow fission. Suppose that the subject's φ-features are copied from T to Aux in (14). Now fission can take place in two places: Aux or T. Two further proposals can now explain OCL-for-SCL. First, suppose that the varieties showing OCL-for-SCL apply fission to the φ-features copied to Aux, while those allowing SCL-OCL-(Aux)-V do it at the T-level. Second, let us introduce the following condition on morphological fusion (fusion is the opposite of fission, the morphological operation of combining feature bundles; see again Halle & Marantz 1993):

(15)   Two feature bundles F_1 and F_2 will fuse **if and only if** they are contained in X, where X is a minimal category containing F_1 and F_2 and there is no minimal category Y such that Y is a minimal category containing X and F_1 and F_2.

Fusion of object and subject φ-features applies to give inverse agreement systems (see Bejar & Rezac 2009; Bárány 2015), e.g. in Hungarian. Hungarian shows agreement with definite direct objects, including 3rd-person pronouns, but does not allow object agreement with a 1st-person pronominal object and a special form appears where the subject is 1SG and the object 2SG:

(16)   a.   Lát-sz   engem.
            See-2SG.SUBJ I.ACC
            'You see me.'

       b.   Lát   engem.
            See-3SG.SUBJ I.ACC
            'He sees me.'

   c.   Lát-lak   téged.
         See-1 > 2 you.SG.ACC
         'I see you.'

A crucial step in Bárány's analysis of this phenomenon is that "T and v undergo fusion … they form a complex head which can be spelled out as a portmanteau morpheme at Vocabulary Insertion" (2015: 10), i.e., morphophonologically realised as a single portmanteau morpheme.

    The Romance languages typically lack the relevant vocabulary items to realise inverse/fused agreement, therefore, in (14), the SCL cannot be phonologically realised under Aux. In this situation, there are three options available. One is to realise the SCL in a structurally higher position than the Aux. This gives rise to a vocalic clitic (as we saw in the Cairese and Fara Novarese examples above, where the SCL is in the left-periphery of the clause to the left of and structurally higher than Aux). A second option is to not realise the SCL at all; no information is lost since the subject's φ-features are present on Aux anyway. A third option is to realise the OCL in a different, lower position, in enclisis to the participle. The relevant structures are shown in (17):

(17)   a.   *$[_T$ $[_{Aux}$ SCL $[_{Aux}$ OCL AUX]] T]
       b. $[_T$ SCL $[_{Aux}$ $[_{Aux}$ OCL AUX] T]
       c.   SCL $[_T$ $[_{Aux}$ $[_{Aux}$ OCL AUX] T]
       d. $[_T$ $[_{Aux}$ $[_{Aux}$ SCL AUX] T]

The structure in (17a), where both clitics attach to Aux, is ruled by the condition on morphological fusion in (15), which requires fusion of the features (this gives rise to inverse marking in languages like Hungarian) and the lack of 'clitic fusion' in Romance. The structure in (17b) is found in dialects not showing OCL-for-SCL; note that SCL attaches to T, not to Aux. (17c) shows the Cairese/Fara Novarese structure where the SCL is realised higher up and further to the left of the T-Aux complex (as a vocalic clitic in these varieties). Finally, (17d) shows the structure of T where the OCL is realised enclitic to participle and so is outside T. Let us now consider this last operation in more detail.

3.2   Enclisis to the past participle

The structure for enclisis to the participle is shown in (18):

(18)   .. Aux $[_{PartP}$ V + Prt $[_{vP}$ $[_v$ OCL + v]] $[_{VP}$ (V) (φ)]]]

Here the verb moves over v to Prt and OCL is enclitic to Prt (this analysis is based on Kayne's 1991 analysis of enclisis to infinitives in languages like Standard Spanish or Italian). As we implied in the previous section, Aux lacks 'object' φ-features; its

φ-feature requirement is satisfied by the SCL alone, thanks to fission and the condition on morphological fusion in (15).

A prediction of the analysis of OCL-for-SCL given in the previous section is that varieties with OCL-enclisis to participles always have a proclitic SCL on Aux. This is certainly true for the Piedmontese varieties discussed in Manzini & Savoia (2005, II: 440–449), as the following examples illustrate:

(19) a. l   ε   cama -me   (Boccioleto)
          3SG.SCL he.IS called = 1SG.ACC
          'He has called me.'

      b. i   ŋ   cama -lu   (Cravagliana)
          3PL.SCL= they.ARE called = 3SGM.ACC
          'They have called him.'

      c. al   a   tʃama -me   (Piverone)
          3SG.SCL= has called = 1SG.ACC

      d. al   a   tʃama -mi   (Moncalvo)
          3SG.SCL= has called = 1SG.ACC
          'He has called me.'

Roberts (2016) suggests that the Romance languages show clitic-climbing in the following three sets of environments:

(20) (i)   complements of restructuring verbs, causatives, and auxiliaries
            (Standard Italian, Spanish, European Portuguese, central-southern Italo-
            Romance, Sardinian, Catalan);
      (ii) complements of causatives and auxiliaries (French, some northern Italo-
            Romance);
      (iii) complements of auxiliaries (Piedmontese and Valdôtain varieties).

Roberts suggests that this could be an implicational scale, although Valdôtain, which clearly has enclisis to past participles and tends to show OCL-for-SCL (see above and Chenal 1986: 384), shows obligatory clitic climbing (at least of the transitive causee) in causatives (Chenal 1986: 535):

(21) a. L' emochon m'   at fé   crié.   (Vaô.)
          the emotion 1SG.ACC= has made.PTCP shout.INF
          'The emotion made me shout.'

      b. Ci travail me   féit   chouére.   (Vaô.)
          this work 1SG.ACC= makes sweat.INF
          'This work makes me sweat.'

      c. Cen m'   at fé   penser   a tè.   (Vaô.)
          that 1SG.ACC= has made.PTCP think.INF to you
          'That made me think of you.'

Also, Parry (1995) appears to suggest that clitic climbing is obligatory in Piedmontese causatives (although Duberti (n.d.) indicates that clitic climbing is not obligatory in the Piedmontese of Mondovì). It is thus possible that the option in (25iii) may not be a related to transparency but instead the consequence of fission applying at the Aux level rather than the T-level in (14).

This suggests the implicational scale 'complement to restructuring verb > complement to auxiliary > complement to causative'. We can describe this as degrees of transparency, which may translate into structural 'size'. Causative complements are the smallest and therefore the most transparent, complements to auxiliaries are intermediate and complements to restructuring verbs largest and therefore least transparent. The implicational scale thus says that if clitic-climbing is allowed from a relatively 'large' complement, it will be allowed from all the smaller ones. Finite clauses are too large to ever allow it.

4.   Conclusion

In the previous sections, I have illustrated and analysed the intricate phenomenon of OCL-for-SCL. This phenomenon, which at first sight appears highly unusual, reflects operations of morphophonological realisation such as fission and fusion, amply attested elsewhere in the world's languages.

Acknowledgements

References

Bárány, András. 2015. Differential Object Marking in Hungarian and the Morphosyntax of Case and Agreement. PhD dissertation, University of Cambridge.

Bejar, Susana & Milan, Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40(1): 35–73. https://doi.org/10.1162/ling.2009.40.1.35

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

Chenal, Aimé. 1986. *Le franco-provençal valdôtain: Morphologie et syntaxe*. Aosta: Musumeci Editeur.

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

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

Déchaine, Rose-Marie & Wiltschko, Martina. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409–442. https://doi.org/10.1162/002438902760168554

Duberti, Nicola. n.d. Causativi di confine. PowerPoint presentation. <https://www.academia. edu/4089533/Borders_causative_constructions>

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, Kenneth Hale & Samual Jay Keyser (eds), 111–176. Cambridge MA: The MIT Press.

Kayne, Richard. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22: 647–686.

Manzini, M. Rita. & Savoia, Leonardo M. 2005. *I dialetti italiani e romanci: Morfosintassi generativa*, 3 Vols. Alessandria: Edizioni dell'Orso.

Muysken, Pieter. 1982. Parametrizing the notion 'Head'. *Journal of Linguistic Research* 2: 57–75.

Parry, Mair. 1994. Posizione dei clitici complemento nelle costruzione verbali perifrastiche del piemontese. *VIII Rëscontr antërnassional dë Studi an sla Lenga e la Literatura Piemontèisa*, 247–259.

Poletto, Cecilia. 2000. *The Higher Functional Field in the Northern Italian Dialects*. Oxford: OUP.

Poletto, Cecilia & Christina, Tortora. 2016. Subject clitics. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds), 772–785. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677108.003.0047

Roberts, Ian. 1993. The nature of subject clitics in Franco-Provençal Valdôtain. In *Dialects of Italy*, Adriana Belletti (ed.), 319–353. Turin: Rosenberg & Sellier.

Roberts, Ian. 2010. *Agreement and Head-Movement. Clitics, Incorporation and Defective Goals*. Cambridge MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262014304.001.0001

Roberts, Ian. 2014. Subject clitics and macroparameters. In *Diachrony and Dialects: Grammatical Change in the Dialects of Italy*, Paola Benincà, Adam Ledgeway & Nigel Vincent (eds), 177–201. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198701781.003.0009

Roberts, Ian. 2016. Object clitics. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds), 786–800. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677108.003.0048

Sound pattern and syntactic structure

Are Sardinian vocatives perfectly regular?

Franck Floricic and Lucia Molinu

University of Paris III / University of Toulouse-Jean Jaurès

The aim of this contribution is to honor Leonardo Savoia with some reflections concerning a topic which has received little attention until recent years: the *case* of Vocatives. If according to D'Alessandro & Van Oostendorp (2016: 78) Vocatives are perfectly regular, typological data show on the other hand that in many languages this is not the case. More specifically, we present and discuss the question of the shape of the Vocatives found in Logudorese Sardinian. We shall bring new data that confirm the view expressed among others in the seminal work of Uspensky & Zhivov (1977) and Floricic (2002, 2011), according to which Vocatives are in many respects *exceptional*.

Keywords: Vocatives, Sardinian, truncation, appeal plane

Introduction

The aim of this contribution is to honor our friend Leonardo Savoia with some reflections concerning a topic which has received little attention until recent years: the *case* of Vocatives. The postulate of such a silence probably lies on the assumption that in a syntax-oriented approach to linguistic structure, there was almost nothing to say about it. Furthermore, if according to D'Alessandro & van Oostendorp (2016: 78) Vocatives are perfectly regular, typological data show on the other hand that in many languages this is not the case. Our aim is to present and discuss the question of the shape of the Vocatives found in Logudorese Sardinian. We shall bring new data that confirm the view expressed among others in the seminal work of Uspensky & Zhivov (1977) and Floricic (2002, 2011), according to which Vocatives are in many respects *exceptional*.

1.   On Vocatives in Sardinian

Refering among others to Vanrell & Cabré (2011), D'Alessandro & Van Oostendorp (2016: 63) observe that "Central and Southern Italian dialects, Sardinian Catalan, Corsican, Sardinian (Alber 2010), as well as regional southern Italian (...) make use of a different reduction strategy whereby the noun is truncated right after the stressed vowel". It then follows that truncation is not a prosodic constituent, but it applies to a phonological word which can be made of various (syntactic) words.

According to the authors, when expressed by a full DP, Vocative truncation would not usually take place. This is questionable: many vocative expressions may take as input a syntactic string. This is the case of example (1a) in Logudorese Sardinian (dialect of Buddusò), where the post-nominal possessive determiner appears in its truncated form: the sequence 'vidzu 'me 'my son' clearly derives from 'vidzu 'meu. The same analysis holds in (1b), where the proper name ni'ɣɔla is modified by the same truncated possessive, thus producing a truncated Vocative in all respects parallel to the other Vocatives found in the language:

(1) a. 'vidzu 'me, 'mi
         'My   son, look!'
     b. ni'ɣɔla 'me
         'My   Nicola!'

By the way it will be noted that truncation also applies to the Imperative 'mi obtained from 'mira 'look!', hence producing an exceptional phonological pattern in Sardinian, not a *regular* one (cf. as well the monosyllabic Imperatives discussed in Floricic & Molinu 2012). Other examples of the same truncation process in DPs are given in (2), taken from Floricic (2002: 165):

(2) a. zu 'ma (< su 'mastru)
         'The master!'
     b. 'tiu vran'tsi (< 'tiu vran'tsisku)
         'Uncle Franziscu!'
     c. zu ðut'tɔ (< su ðut'tɔrɛ)
         'The doctor!'
     d. zu ðut'tore'me (< su ðut'tɔrɛ 'meu).
         'My doctor!'
     c. su bo'no (< su bo'ninɛ)
         'The guy!'

D'Alessandro & van Oostendorp (2016: 67) would probably exclude that we are dealing here with *true Vocatives*. As a matter of fact they assume the distinction made by several scholars who set apart true Vocatives and fake Vocatives: "true vocatives

are deictic, while fake ones are full DPs, associated with a predicational meaning". Following Espinal (2013: 109) they assume the existence of a Voc(ative) head in the left periphery, endowed with a deictic feature [+DX] referring to the addressee.

Needless to say, such a definition – couched within a minimalist framework – is highly restrictive. In this framework, *true Vocatives* are always deictic in the sense that they have to move to the Voc head to 'check' the [+DX] feature. According to Espinal, vocative particles should be expected with *true Vocatives*, while they would be ruled out with *fake Vocatives*: given that the vocative particle specifies the Voc head, it would be incompatible with a full DP. It must be observed, however, that examples such as (3a)-(b) are perfectly acceptable, despite the fact that the vocative particle precedes a full DP:

(3)　a.　ɔ zu 'ma (< su 'mastru)
　　　　　'O the master!'
　　b.　ɔ 'tiu vran'tsi (< 'tiu vran'tsisku)
　　　　　'O uncle Franziscu!'
　　c.　fran'tsi / ɔ vran'tsi (< fran'tsisku)
　　　　　'O Franziscu!'
　　d.　pje'ri / ɔ βje'ri (< pje'rina)
　　　　　'O Pierina!'
　　e.　'pe / ɔ 'βe (< pedru)
　　　　　'O Pedru!'

It is hard to assume that the expressions (2a)-(2b) and (3a)-(3b) would be non-deictic: they have an identificational *effect* (even though their *function* is to catch / maintain the addressee's attention) as do the other Vocatives in (3c)-(3e). Interestingly, the truncated proper name in the Vocative can be repeated, with the second occurrence preceded by the vocative particle [ɔ], especially when the individual called for does not answer immediately.[1]

In other words and to sum up, a vocative expression identifies the entity (mostly human but not only) whose attention is called upon in a given situation (and in a given spatio-temporal setting). In this sense, as a "gesture translated into sound" (Schleicher 1862: 162), the Vocative has much in common with pointing and gestures in general than with conceptual reference. But if the identification process is direct in the case of proper names, it may be said to be *mediated* in the case of a DP like zu 'ma 'the master!'. Needless to say, the pragmatic efficiency of Vocatives relies fundamentally on intonation / prosody. It must be borne in mind as well that the formal properties of Vocatives heavily depend on such parameters as politeness strategies, social rank, sex (male and female contrast), familiarity / formality, emphasis (insistence), irritation, distance, visibility, etc. Recall

---

1. Observe that the reversed order is perfectly possible as well: ɔ βje'ri / pje'ri.(< pje'rina).

for example that in a language like Ket (Siberian), the Vocative has a dynamically stressed final vowel rather than a regular tonal contour (Vajda 2007: 1281–1282), and the feminine singular nouns distinguish proximal Vocatives (cf. *amá* 'Mother!' [near or visible]), which are expressed with a stressed [a] suffix, from distal ones (*amə* Mother!' [distant or out of sight]), expressed with stressed [ʌ]. Distal and proximal Vocatives are also reported in Hualapai, where the suffix *-é* is used in the singular (*−yé* in the plural (Watahomigie, Bender & Yamamoto 1982: 71–74)) when the addressee is near the speaker, while the suffix *-ó* / *-wó* is required when the addressee is out of sight (*−wó* is added when the noun ends in a long vowel, while *-ó* is required in all the other environments).

Abnormal final vowel lengthening answers the same purpose: in Yupik (Central Alaskan), the doubling of the final vowel of absolutives characterizes a vocative form used in addressing persons at some distance or in an exaggerated way (Woodbury 1987: 726; Miyaoka 2012: 143, 175, 859, etc.)[2]:

(4)   Absolutive   Distal Vocative
   a.   [ukút] 'these here'   [úk.:ú:t] '(Hey,) you (p1.) here!'
   b.   [an.ŋaqɬíq] 'eldest brother'   [án.ŋáq.ɬí:q] '(Hey,) eldest brother!'
   c.   [á:.ná] 'mother'   [á:.ná:] '(Hey,) Mother!'
   d.   [nupíx.:áq] '(personal name)'   [nupíx.:á:q] '(Hey,) N.!'
   e.   [qayáq] 'kayak'   [qáy.:á:q] '(Hey,) kayak(er)!'
   f.   [qəŋáx̣.palək] 'big-nosed'   [qəŋáx̣.pál.:í:k] '(Hey,) big-nosed!'

As illustrated in (4f), in the Vocative the final vowel /ɨ/ is doubled into /ii/ (cf. Miyaoka 2012: 143). The interesting aspect to be pointed out is that this anomalous feature iconically answers the purpose of attracting the hearer's attention (cf. Jacobsen 1994: 34). Word final long vowels are also attested in the Vocatives of Nivkh (cf. Nedjalkov & Otaina 2013: 55), where the final vowel of proper names or nouns is lengthened or made more open (cf. *ōla-gu* 'children' > *ōla-gū* / *ōla-gō*!) – with nouns ending in a consonant, the suffix *-a* is added (cf. *ətək* 'father' > 'father!').[3] It

---

2. Interestingly, final vowel doubling in Vocatives is avoided when addressing to an elder or respected person (cf. Miyaoka 2012: 188).

3. The situation however is a little bit more complex in so far as in Nivkh, the stem of proper names or kinship terms to which the vocative marker attaches may be shortened (I am indebted to Ekaterina Gruzdeva for the following data):

(5)   Base form   Vocative   Gloss
   a.   əmk   əm-a: 'mother!
        (East Sakhalin dialect)
   b.   aṭmam   aṭm-a: 'grand-mother!'
        (East Sakhalin dialect)

is attested as well in Malayalam (cf. Asher & Kumari 2013: 223–224), where "Male names in final -*an* replace it by -*ā*, -*m* ending female names double the final -*m*, all vowel-final names lengthen the vowel and other consonantal-final words add -*ē* in address, e.g. *mādhavan*: *mādhav-ā*, *kamala-m*: *kamala-mm*, *makka.l* 'children (one's own)': *makka.l-ē*, *rā.ṇi*: *rā.ṇ-ī* (proper noun)" (Krishnamurti 2003: 458).

Word final long vowels are reported in the Vocative of Southern Sierra Miwok (Jacobsen 1994: 34–35) and in Chipewyan (where the Vocative is formed either by lengthening a final vowel and imposing a falling tone on it; or by adding a suffix -*î* which replaces a final vowel). In the same way, Hidatsa (Siouan) is a pitch-accent language where pitch is used in the formation of the Vocatives: as illustrated in (6), the vocative forms of this language have a falling pitch on the last syllable and the last vowel is lengthened if it is short (cf. Davis & Tsujimura 2014: 216):

(6)   Base form Vocative Gloss
    a.   marisá   marisáà 'my son!'
    b.   masáàwi   masaawîi 'my aunt (father's sister)!'
    c.   magúù   magúù   'my grandmother!'
    d.   masígisa masigisáà 'my brother-in-law (women's brother-in-law)!'

In Manambu, vocative forms are marked with a *ə* plus an off-glide *y*, or with *ay* / *a*. The final syllable undergoes lengthening, as illustrated by the Vocative *Walupa:y!* of the name *Walup*. Aikhenvald (2009: 44) observes that when *ə* undergoes lengthening in a vocative form, it is pronounced as *e:* (cf. the Vocative *Maliyée:(y)* of the name *Máli*, where stress moves to the last syllable).[4] Here too, we are crucially dealing with abnormal / unusual phonetic patterns. Needless to say, many other examples could be adduced to illustrate the exceptional properties of Vocatives.

From the Sardinian examples presented in (1–3), it thus follows that a) truncation applies regardless of the fact that the Vocative is represented by a (simple) proper name or a full DP; b) the Vocative is formed deleting all the material

---

    c.   aki   ak-a: 'elder.brother!'
          (East Sakhalin dialect)
    d.   aṭik   aṭ-a: 'younger-brother!'
          (Amur dialect)
    e.   ṭenduk   ṭen-a: 'Tenduk!'
          (Amur dialect)

Observe that in his presentation of written Mongolian, Poppe (2006: 138–139) holds the view that there is no vocative suffix in this language and that the nouns used in direct address appear in the nominative plus the interjection *a:* added to the noun.

4. As Aikhenvald (2009: 59) points out, "Vocative intonation involves slight rise on the last syllable of the vocative and simultaneous lengthening of vowel in the last syllable of the vocative forms (restricted to kinship terms and personal names)". See the case of the name *Maliye::: ↗*.

following the stressed syllable; c) if the stressed syllable is represented by a CVC syllable in the input, the Vocatives only retains the CV string; d) lenition applies between the vocative particle and the initial consonant of the following expression, thus recalling a phenomenon widely attested in Celtic languages, but examples (2a)-(2b) show that lenition may also apply in absolute initial position; e) for emphasis reasons (insistence, distance, etc.), the final vowel of the truncated Vocative may be lengthened as it is in Manambu (cf. the notion of *plutivokativ* discussed among others by Loewe (1923), Nehring (1933), Bravmann (1935)): in this case, a particular intensity on the last syllable can be coupled with overlength of the final vowel (cf. Stankiewicz 1964: 248); f) when the vocative is preceded by the particle ɔ, this particle bears the high pitch that we can otherwise find on the first syllable of the (bare) vocative form.[5] From this point of view, it is not clear whether we should follow d'Alessandro & Van Oostendorp (2016: 77) when they state that "High pitch is impossible to realize on an unstressed or reduced vowel".

The point to be stressed is that the beginning of the vocative expression tends to be maximized or strengthened. We are dealing here with calling forms belonging to the *appeal plane* of language (cf. Bühler 2011: 35) and the phonetic distortions we may find in this kind of forms is tightly linked to this calling function (cf. Kuryłowicz 1975; Isačenko 1964: 89, a.o.). But not all the vocative forms are of the same kind and some of them may be more or less integrated into the syntactic structure of the sentence. Let us add that the accentual pattern illustrated in (3) is exceptional in Sardinian: it is only attested with truncated (interjectional) Imperatives (cf. Floricic 2011; Floricic & Molinu 2012), primary interjections, nursery words or with (non-integrated) loanwords. By the way, another oddity of these forms is exemplified in (3e) and in (7a-h):

(7)   a.   anˈtɔ (< anˈtɔna).
         'Anto!'
      b.   anˈto (< anˈtoni).
         'Anto!'

---

5. In various languages, the same kind of particle may follow the vocative. In Bislama, the particle *o* may follow the name in the Vocative: in this case (cf. *Kalteri o!* 'Hey, Kalteri!'), Crowley (2004: 32) assumes that it is "pronounced with a high intonation and lengthened vowel". The same observation holds for Harari, where >*o* is added at the end of the name: *ahlače-o* 'o my friends!'. In Geʿez, the same element can appear either at the left (cf. >*o-gabr* 'o servant') or at the right of the name (cf. *bə* > *əsit-o* 'o woman'), or both (cf. >*o-bə* > *əsit-o* 'o woman') (cf. Leslau 1945: 73 and 2006: 1). In Albanian too, according to Camaj (1984: 36) the "exclamatory particle" *o* may be preposed or postposed to the proper name: *O Artan!* / *Artan-o!* 'Artan!'; *o bir!* / *bir-o!* 'oh son!' (Newmark et al. 1982: 134–135 exclude the Vocative from the case inventory on the grounds that it is always undistinguishable from the nominative both in the definite and in the indefinite form).

    c.   antɔ'nɛ (< antɔ'neɖɖa).
       'Antonedda!'
    d.   anto'ne (< anto'neɖɖu).
       'Antoneddu!'
    e.   mau̯'re (< mau̯'reɖɖu).
       Maureddu!
    f.   mau̯'rɛ (< mau̯'rɛɖɖa).
       Mauredda!
    g.   bus'tja (< bus'tjanu / bus'tjana).
       'Bustià!'
    h.   'ʒwa (<'ʒwannɛ / 'ʒwanna).[6]
       'Ciuà!'

Leaving aside the case of final stress which is once again exceptional, it can be observed that in the pairs (7a)-(7b), (7c)-(7d) and (7e)-(7f) the gender opposition is marked by the contrast between mid-high and mid-low vowels: in truncated forms like an'to or anto'ne, the final (stressed) mid-high vowel is due to an harmony process whereby the final high vowels [i], [u] of the base trigger raising of the preceding mid-low vowel. Such examples nicely illustrate the fact that phonological processes are ordered: metaphony occurs first, and truncation applies to the output of such operation (the same observation holds for the nice example given in (2e) (i.e. su bo'no < su bo'nominɛ), where the high front vowel [i] triggers the raising of the preceding back vowel).

    Even though the phonology of Sardinian rules out mid-high final vowels, they are perfectly attested in interjections and Vocatives, where they mark the contrast between masculine and feminine proper names; but the contrast between mid-high and mid-low vowels, limited to this marginal sub-domains, does not entail recognition of their phonemic distinctiveness. The case of bus'tja (< bus'tjanu / bus'tjana) and 'ʒwa (<'ʒwannɛ / 'ʒwanna) in (7g)-(7h) interestingly shows that with other proper names, truncation may lead to neutralization of gender oppositions (cf. Stankiewicz 1961, 1964).

2.   The syntax of Vocatives

As far as syntax is concerned, D'Alessandro & van Oostendorp (2016) point out that Vocative is not a structural case but a left-peripheral element, although it is

---

6. In our view, it is hard to discriminate between hiatus (i.e. ʒu'annɛ) and diphthong (i.e. 'ʒwannɛ); it is clear that allegro speech (vs. careful and slow elocution) may favor one or the other.

a case form in the European grammatical tradition. However it must not be forgotten that linguistic processes and categories are not fixed and eternal platonic objects: they are dynamic and Kuryłowicz (1975: 54) points out that "The vocative continues an old nominative forced out of the cadre of the declension, thus vocative plural O. Irish *firu!* 'o men!' from nominative plural **u̯irōs* ousted by the new nominative **u̯iroi (O.* Irish *fir)*". The evolution of languages thus shows that the vocative case can be lost and that new forms of vocative may be created with specific exponents or with specific morphological operations (cf. Floricic 2011; Andersen 2012; Holvoet 2012; Paulis 2013; Janda to appear) – non solidarity is one of the main features of the Vocative case. Furthermore, as Kuryłowicz (1975: 136) points out, "Beaucoup de nominatifs sing, se sont fixés sous la forme de l'ancien vocatif, cf. les noms propres béotiens Βουκάττες, Μέννει etc., lat. *Iuppiter (<*dieu pater),* s. cr. *Vukašine kralju, kraljeviću Marko,* pol. *Jasiu, Franiu* etc.". Hence some kind of mismatch between morphological information and syntactic behaviour.

For sure, the Vocative is not a phrase belonging to the thematic grid of a predicate. But this is not a reason to assume that the Vocative is a-syntactic or that it lies outside syntax. It was pointed out at least since Tesnière (1988) that the Vocative shares some properties with Topics. Both are extrapropositional and both are separated from the rest of the clause by a pause; as illustrated in (8a)-(8b), the proper name 'ʒwannɛ is coindexed with the object clitic preceding the verb of the following clause (the initial preposition in (8b) is the exponent of the DOM strategy found in Sardinian):

(8) a. 'ʒwannɛ / a tti 'βɔttɔ aʒu'arɛ.
        'Ciuanne, can I help you?'
    b. a t'ʧwannɛ / 'l appo ɣon'nottu ɣust is'tiu.
        'Ciuanne, I knew him this summer'.
    c. (ɔ) 'ʒwa / a tti 'βɔttɔ aʒu'arɛ.
        'Ciuanne, can I help you?'
    d. *a t'ʧwa / 'l appo ɣon'nottu ɣust is'tiu.
        'Ciuanne, I knew him this summer'.
    e. ti 'βɔttɔ aʒu'are /??(ɔ) 'ʒwa.
        'Can I help you, Ciuanne'.

The examples (8c)-(8d) however show that the parallelism between the Vocative and the Topic cannot be pushed too far: the truncated form of the proper name 'ʧwannɛ cannot be used as a Topic (cf. Floricic 2002: 168). And if the proper name can be used as a Vocative in post-clausal position, as is the case of after-thoughts, the calling Vocative preceded by the vocative particle [ɔ] is excluded in this position (cf. (8e)). The point to be stressed is that there is no reason to deny the status of *Vocative* to the truncated proper names in (3) and (8). Leaving aside

coindexation between the Vocative proper name and the subject / object of the following / preceding verb, concord phenomena show that dependent elements can bear the same vocative exponents of the head noun / proper name of the vocative DP. This is particularly clear in languages like Romanian or Georgian, as illustrated in (9a)-(9b) and (9c)-(9d) respectively (cf. Sandfeld 1930: 146; Daniel & Spencer 2009: 633, a.o.):

(9)   a.   stimate domnule!
             'Dear Sir!'
      b.   milostive şi luminate doamne!   (cf. Croitor 2014: 81).
             'Merciful and enlightened God!'
      c.   čem-o k'arg-o.
             my-VOC dear-VOC.
             'My dear'.
      d.   tkven-o ağmatebuleba-v   (cf. Abuladze & Ludden 2013: 28).
             'Your Eminence!'
      e.   p'at'ivtsemulo bat'ono da megobaro   (G. Hewitt, p.c.)
             'Dear Sir and friend'.

In the modern Romanian example (9a), the adjective *stimate* agrees with the articulated vocative noun *domnule*; in the Old Romanian example (9b) the two conjoined adjectives *milostive* and *luminate* agree with the head vocative noun *doamne*. The same observations holds for the Georgian examples in (9c)-(9e): the vocative exponent *-o* is found on both the governing and the dependent element of the DP – *-v* in (9d) is an allomorph of *-o*. And the vocative nouns perfectly admit coordination (cf. (9e)).

3.   Conclusion

Needless to say, the preceding lines cannot do justice of the complex problem raised by the Vocative and the question of casehood. The Sardinian data show that in this language there are formally distinct expressions dedicated to the calling function. The truncation processes which are typically found in vocative formation may produce very short forms resulting from deletion of whatever follows the stressed syllable. At the same time, the very nature of the calling Vocatives and their efficiency rely on some kind of maximization of the beginning of the word, even though in some cases the right edge of the vocative expression may show as well strengthening processes. From this point of view, the phonetic properties of the Vocative may be said to iconically reflect its calling function.

References

Abuladze, Lia & Ludden, Andreas. 2013. The vocative in Georgian. In *Vocative!. Addressing between System and Performance*, Barbara Sonnenhauser & Patrizia Noel Aziz Hanna (eds), 25–42. Berlin: De Gruyter. https://doi.org/10.1515/9783110304176.25

Aikhenvald, Alexandra. 2009. *The Manambu Language East Sepik, Papua New Guinea*. Oxford: OUP.

D'Alessandro, Roberta & van Oostendorp, Marc. 2016. When imperfections are perfect. Prosody, phi-features and deixis in Central and Southern Italian Vocatives. In *Romance Languages and Linguistic Theory 10: Selected Papers from 'Going Romance' 28 Lisbon* [Romance Languages and Linguistic Theory 10], Ernestina Carrilho, Alexandra Fiéis, Maria Lobo & Sandra Pereira (eds) , Maria Lobo & Sandra Pereira (eds), 61–82. Amsterdam: John Benjamins. https://doi.org/10.1075/rllt.10.04dal

Andersen, Henning. 2012. The new Russian vocative: Synchrony, diachrony, typology. *Scando-Slavica* 58(1): 122–167. https://doi.org/10.1080/00806765.2012.669918

Asher, Ronald E. & Kumari, T. C. 2013. *Malayalam*. London: Routledge.

Bravmann, Max. 1935. Des effets de l'accentuation emphatique sur la formation de la langue en sémitique. *Mémoires de la Société de Linguistique de Paris* 23: 329–358.

Bühler, Karl. 2011. [1934]. *Theory of Language: The Representational Function of Language*. Translated by Donald Fraser Goodwin in collaboration with Achim Eschbach. Amsterdam: John Benjamins. https://doi.org/10.1075/z.164

Camaj, Martin. 1984. *Albanian Grammar*. Wiesbaden: Harrassowitz.

Croitor, Blanca. 2014. Vocative markers in Old Romanian. In *Limba română: Diacronie și sincronie în studiul limbii române (I). Actele celui de-al 13-lea Colocviu Internațional al Departamentului de Lingvistică* (București, 13–14 decembrie 2013), Rodica Zafiu, Adina Dragomirescu & Alexandru Nicolae (eds), 77–86. București: Editura Universității din București.

Crowley, Terry. 2004. *Bislama Reference Grammar*. Honolulu HI: University of Hawai'i Press.

Daniel, Michael & Spencer, Andrew. 2009. Vocative: An outlier case. In *The Oxford Handbook of Case*, Andrej Malchukov & Andrew Spencer (eds), 626–634. Oxford: OUP.

Davis, Stuart & Tsujimura, Natsuko. 2014. Non-concatenative derivation: Other processes. In *The Oxford Handbook of Derivational Morphology*, Rochelle Lieber & Pavol Stekauer (eds), 190–218. Oxford: OUP.

Espinal, Maria T. 2013. On the structure of vocatives. In *Vocative!. Addressing between System and Performance*, Barbara Sonnenhauser & Patrizia Noel Aziz Hanna (eds), 109–132. Berlin: De Gruyter. https://doi.org/10.1515/9783110304176.109

Floricic, Franck. 2002. La morphologie du vocatif : l'Exemple du sarde. *Vox Romanica* 61: 151–177.

Floricic, Franck. 2011. Le vocatif et la périphérie du système des cas: Entre archaïsmes et innovations. In *L'évolution grammaticale à travers les langues romanes* [Mémoires de la Société de Linguistique de Paris 19 (Nouvelle Série)], 103–134. Louvain: Peeters.

Floricic, Franck & Molinu, Lucia. 2012. Romance monosyllabic imperatives and markedness. In *Monosyllables. From Phonology to Typology*, Thomas Stolz, Nicole Nau, & Cornelia Stroh (eds), 149–172. Berlin: Akademie Verlag.

Holvoet, Axel. 2012. Vocative agreement in Latvian and the principle of morphology-free syntax. *Baltic Linguistics* 3: 43–64.

Isačenko, Alexandr. 1964. On the conative function of language. In *A Prague School Reader in Linguistics*, Joseph Vachek (ed.), 88–97. Bloomington IN: Indiana University Press.

Jacobsen, William H. 1994. Nootkan vocative vocalism and its implications. In *Sound Symbolism*, Leanne Hinton, Johanna Nichols & John Ohala, (eds), 23–39. Cambridge: CUP.

Janda, Laura. To appear. Name-calling: The Russian 'New Vocative' and its Part of Speech. ms.

Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486876

Kuryłowicz, Jerzy. 1975[1968]. La gémination consonantique dans les noms propres. In *Esquisses Linguistiques* 2: 135–142. Munich: Wilhelm Fink.

Leslau, Wolf. 1945. The influence of Cushitic on the Semitic languages of Ethiopia. A problema of substratum. *Word* 1: 59–82. https://doi.org/10.1080/00437956.1945.11659246

Leslau, Wolf. 2006. *Comparative Dictionary of Geʿez (classical Ethiopic): Geʿez-English, English-Geʿez, with an Index of the Semitic Roots*. Wiesbaden: Otto Harrassowitz.

Loewe, Richard. 1923). Die indogermanische Vokativbetonung. *Zeitschrift für Vergleichende Sprachforschung* 51: 67–108; 161–220.

Miyaoka, Osahito. 2012. *A Grammar of Central Alaskan Yupik*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110278576

Nedjalkov, Vladimir P. & Otaina, Galina A. 2013. *A Syntax of the Nivkh Language. The Amur Dialect* [Studies in Language Companion Series 139], translated from Russian by Emma Š. Geniušienė, edited by Emma Š. Geniušienė & Ekaterina Gruzdeva. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.139

Nehring, Adolf. 1933. Anruf, Ausruf und Anrede. In *Festschrift Theodor Siebs zum 70. Geburtstag 26. August 1932*, Walther Steller (ed.), 95–144. Breslau: M. & H. Marcus.

Newmark, Leonard, Hubbard, Philip L. & Prifti, Peter R. 1982. *Standard Albanian: A Reference Grammar for Students*. Stanford CA: Stanford University Press.

Paulis, Giulio. 2013. Dinamiche linguistiche e sociali a Cagliari nel corso dei secoli. Le alterne fortune di un "vocativo" di lunga durata. In *Repertorio plurilingue e variazione linguistica a Cagliari*, Giulio Paulis, Immacolata Pinto & Ignazio Putzu (eds), 40–81. Milano: Franco Angeli.

Poppe, Nicholas. 2006. *Grammar of written Mongolian*. Wiesbaden: Harassowitz.

Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats* [Société de Linguistique de Paris XXXI]. Paris: Champion.

Schleicher, August. 1862. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*, II: *Formenlehre*. Weimar.

Stankiewicz, Edward. 1961. Grammatical Neutralization in Slavic Expressive Forms. *Word* 17: 128–145. https://doi.org/10.1080/00437956.1961.11659749

Stankiewicz, Edward. 1964. Problems of emotive language. In *Approaches to Semiotics*, Thomas A. Sebeok, Alfred. S. Haye &, Mary K. Bateson (eds), 239–264. The Hague: Mouton.

Tesnière, Lucien. 1988 [1959]. *Eléments de syntaxe structurale*. Paris: Klincksieck.

Uspensky, Boris A. & Zhivov, Viktor M. 1977. Center – Periphery opposition and language universals. *Linguistics*: 196: 5–24.

Vajda, Edward J. 2007. Ket morphology. In *Morphologies of Asia and Africa*, Alan S. Kaye (ed.), 1277–1325. Winona Lake IN: Eisenbrauns.

Vanrell, Maria del Mar & Cabré, Teresa. 2011. Troncamento e intonazione dei vocativi in Italia centromeridionale. In *Contesto comunicativo e variabilità nella produzione e percezione della lingua, Atti el 7 convegno AISV(26–28 January 2011, Lecce)*, Antonio Stella, Barbara Gili-Fivela, Luigia Garrapa & Mirko Grimaldi (eds). Rome: Bulzoni.

Watahomigie, Lucille J., Bender, Jorigine & Yamamoto, Akira Y. 1982. *Hualapai Reference Grammar*. Los Angeles LA: American Indian Studies Center.

Woodbury, Anthony C. 1987. Meaningful phonological processes: A consideration of central Alaskan Yupik Eskimo prosody. *Language* 63: 685–740. https://doi.org/10.2307/415716

CHAPTER 19

Phonological correlates of syntactic structure

The distribution of *raddoppiamento fonosintattico* in Calabrian

Adam Ledgeway

University of Cambridge

The present article considers the syntactic constraints operative on the distribution of a phonological fortition process, *raddoppiamento fonosintattico* 'phonosyntactic doubling', in the Calabrian dialect of Cosenza. It is shown that the relevant locality restrictions are best understood, not in terms of the three core structural configurations Spec-Head, Head-Head and Head-Comp, but in terms of phasal domains, highlighting how different phonological realizations represent the spell-out of deep syntactic differences mapped at the syntax-phonology interface. At the same time, the theoretical assumptions assumed here provide us with the key to understanding some intriguing empirical generalizations about the distribution of Cosentino RF which, in turn, throw new light on some current theoretical assumptions about clause structure and the nature of phases.

Keywords: syntax-phonology interface, phase, locality, *raddoppiamento fonosintattico*

1.   Introduction

Through several studies (Savoia 1987; Savoia 2015: Chapter 6; Rizzi & Savoia 1993; Manzini & Savoia 2016), Leonardo Savoia has shown how the surface distribution of particular phonological phenomena in the dialects of southern Italy such as the propagation of /u/ and restoration of −/a/ are constrained by specific syntactic constraints. Taking inspiration from his work, in what follows we consider within a broadly minimalist framework how in the southern Italian dialect of Cosenza syntactic constraints determine the distribution of another phonological process, namely *raddoppiamento fonosintattico* (RF) 'phonosyntactic doubling' (Rohlfs 1969: 235–238; Loporcaro 1988, 1997; Vincent 1988; Fanciullo 1997; Ledgeway 2016a: 214; Sampson 2016: 675–676), an initial-consonant lengthening process

which originates as an external sandhi assimilation triggered by a small class of words that historically ended in a final consonant, e.g. Cosentino *cchiù* (<PLUS) 'more' + *granne* 'tall' > *cchiù ggranne* 'taller'.[1] While superficial phonological adjacency appears sufficient for RF to apply, it has been known since at least D'Ovidio (1874: 179) that the process in southern Italy is also constrained by structural conditions, a situation insightfully captured by Fanciullo's (1986: 88) observation that RF "occurs only if the lexeme which causes it constitutes, together with the item it acts on, a *minimal phrase* – a kind of hierarchically superior word" (cf. also Andalò 1991; Loporcaro 1997: 49; Savoia 2015: 436–441). In short, Fanciullo's intuition is to link the phonological licensing of RF to constituency, hence its absence in *granne cchiù 'i* (*ddi < DE) *mia* 'tall more of (= than) me', where post-adjectival *cchiù* continues to modify, and hence form an immediate constituent with, the adjective *granne*, but not the linearly adjacent prepositional complementizer (DE > *di*) *'i* introducing the standard of comparison. Developing this idea further, Ledgeway (2009, 46–47) shows how in Neapolitan the relevant locality constraint on the application of RF encapsulated in Fanciullo's *minimal phrase* can be broken down into the three core structural configurations of Spec-Head, Head–Head and Head-Comp(lement), with RF failing to apply in linearizations falling outside of these configurations.

However, such structural conditions are not sufficient to capture the distribution of RF in all southern varieties, as revealed by Biberauer & D'Alessandro's (2006) analysis of eastern Abruzzese (see now also D'Alessandro & Scheer 2015), where the application of RF in otherwise identical Head-Comp linearizations with passive/active auxiliary BE (e.g. SUM>*so* 'I am') and perfective participle (e.g. *viste* 'seen') is only licensed under the passive interpretation (viz. *so vviste* 'I am seen'), but not its active variant (viz. *so viste* 'I have seen').[2] Adopting a cyclical approach to Spell-Out, they derive this contrast from the variable phasal status of the participial *v*P which, on standard assumptions, only constitutes a phase when active. Consequently, in the active Head-Comp sequence consonantal lengthening fails to obtain because active auxiliary and participle are sent to PF in separate cycles, whereas in the passive string both passive auxiliary and participle are contained within the same higher CP phase and sent to PF together in the same cycle where

---

1. For expository convenience, we indicate consonantal lengthening orthographically with an initial double consonant in bold-type, while acknowledging that the relevant contrast is not simply one of length but, rather, one of fortition involving in some cases a change of manner and place of articulation or the restoration of an underlying consonant (cf. Savoia 2015: 415–416; Ledgeway 2016b: 252–253), e.g. *cchiù* 'more' + *jancu* 'white' > *cchiù gghjancu* 'whiter' (/j/ > [ɟː]); *cchiù* 'more' + *a* 'the' (*carne* 'meat') > *ccchiù ra carne* 'plus the meat' (/Ø/ > [r / ɾ]).

2. See also the discussion in § 3.

the auxiliary can license consonantal lengthening of the adjacent passive participle. In a similar vein, Silvestri (2007, 2014) shows how in the northern Calabrian dialect of Verbicaro the distribution of RF is also restricted by phasal domains, highlighting, in particular, how the presence or absence of RF in V_{3SG} + XP sequences, for example, maps onto different pragmatic interpretations and argument structures.

In the wake of these latter studies, the present article, which forms part of a larger and more detailed investigation (cf. Ledgeway in prep.) of the syntactic constraints on the application of RF in Cosentino, undertakes a brief overview of the structural conditions regulating Cosentino RF. On the one hand, the data highlight the advantages of not interpreting locality narrowly in terms of the three core configurations noted above in Ledgeway (2009), but more broadly in terms of phasal domains, showing how different phonological realizations represent the spell-out of deep syntactic differences mapped at the interface between narrow syntax and PF. On the other, the theoretical assumptions adopted here provide us with the key to understanding some intriguing empirical generalizations about the distribution of Cosentino RF which, in turn, throw new light on some current theoretical ideas about clause structure and the nature of phases.

2.   Cosentino RF: An overview

2.1   RF triggers

An exhaustive list of RF-triggers in Cosentino is given in (a)-(h):

a. Conjunctions: (ET>) *e* 'and', (AC>) *a* 'and' (limited to now lexicalized *diciassette/diciannove* 'seventeen/eighteen' < DECEM AC SEPTEM/NOUEM);
b. Quantifiers: (OMNES>) *ogni* 'each', (TRES>) *tri* 'three';
c. Wh-phrases: (QUID>) *cchi* 'what?/!'
d. (Religious) titles: (MATER>) *matre* 'mother', (PATER>) *patre* 'father';
e. Negators: (NON>) *no/u(n)* 'not', (NEC>) *né* 'neither, nor';
f. Prepositions: (AD>) *a* 'to, at', (CUM>) *ccu* 'with', (PER>) *ppi* 'for';
g. Adverbs: (*ECCU-HAC>) *ccà* 'here', (ILLAC>) *ddà* 'there', (PLUS>) *cchiù* 'more', (*ECCU-SIC>) *accussì* 'so, thus';
h. Verbs: (SIS>) *si* 'you.SG are', (SUN(T)>) *su* 'they are', and all 3SG finite verb forms (< −T);

Although the inventory proves extremely limited, with RF licensed after just a handful of members of individual word classes (cf. *cchi* 'what?' [+RF] vs *picchí* 'why?' [−RF] < (PER)QUID), the incidence of RF is anything but rare, in that the

relevant triggers represent some of the most frequently recurring grammatical items in the dialect. Moreover, in the case of (h) the effects of original underlying final -T in all 3sG finite verb forms guarantees that the actual number and frequency of lexical items able to trigger RF are hardly negligible (Loporcaro 1997: 114–117), especially since RF has been analogically extended to the 1sG in the imperfect indicative and subjunctive and the conditional where 1sG and 3sG are today homophonous (e.g. CLAMABA-M/–T > *chiamava* 'I/(s)he called').[3]

2.2 Defining locality

2.2.1 *Core configurations: Spec-head, head–head and head-comp*
To understand the structural domains in which RF is found, as an initial generalization we note that it is readily licensed in all three of the core local configurations illustrated in (1a-c).[4]

(1) a. [$_{AdvP}$ [$_{Spec}$ cchiù] **cch**ianu] (Spec-Head)
         'more slowly'
    b. [$_{TP}$ [$_{T'}$ [$_{T°}$ [$_{Neg}$ un] [$_T$ [$_{Cl}$ **tt**i]   [$_T$ capisciu]]]]] (Head–Head)
                    not =   you.SG.ACC =   understand.1sG
       'I don't understand you'.
    c. [$_{\&P}$ [$_{Spec}$ Maria] [$_{\&'}$ e   [$_{DP}$ **nn**ua]]] (Head-Comp)
              Maria   and   we
       'Maria and us'.

In (1a) the degree adverb *cchiù* enters into a Spec-Head relation with the manner adverb *chianu* lexicalizing its specifier, a sufficiently local position from which to license RF on the latter. In (1b) the object clitic *ti* and the clitic negator *un* both raise to left-adjoin into the finite verb, itself raised to a functional head within the T-domain, placing the negator and clitic in a Head–Head relation as part of a complex head in which the former triggers RF on the latter (and often giving rise to the further assimilated form *u' tti*). Finally, in (1c) the coordinating conjunction selects a pronominal complement, a D head with null NP complement, producing a canonical Head-Comp configuration in which the head triggers initial-consonantal lengthening of its pronominal complement.

---

3. In a small number of cases (cf. *ha* 'have', *sta* 'stand, be', *fa* 'do'), the presence of RF also serves to distinguish the 3sG from otherwise homophonous 2sG forms, e.g. *ha (*p)pagatu* 'you have paid' vs *ha **pp**agatu* 's/he has paid'.

4. For expository clarity, structural representations have been considerably simplified, omitting, for example, lower copies and all projections not immediately relevant to the discussion.

However, these three configurations must be understood more broadly to also accommodate structures like (2a-c) in which doubler and doublee (henceforth Word$_1$ and Word$_2$), although linearly contiguous, are not immediately adjacent in structural terms.

(2)   a.   [$_{DP}$ [$_{Spec}$ cchiù] [$_{D'}$ D°– [$_{NP}$ ccasu]] (Spec-Head)
           'more   cheese'

      b.   [$_{ForceP}$ [$_{Force'}$ Fa-   [$_{FinP}$ [$_{Fin'}$ [$_{Fin°}$ [$_{Cl}$
                                  do.IMP.2SG   llu]   [$_{Fin}$ ~~fa~~]
        [$_{v\text{-}VP}$ ~~lu fa~~!]]]]]] (Head–Head)   =it.ACC.MSG
        'Do it!'

      c.   [$_{\&P}$ [$_{Spec}$ pane] [$_{\&'}$ e   [$_{DP}$ [$_{D'}$ D°__ [$_{NP}$ ccasu]]]]] (Head-Comp)
           'bread   and   cheese'

In (2a) *cchiù* is now merged in SpecDP, but the head of the projection remains empty leaving the adverb to modify the empty head's NP complement instantiated by the non-projecting nominal head *casu*. We thus obtain a Spec-Head configuration, albeit one in which between the adverb and the nominal head there intervenes a non-lexicalized D° head. Similarly, in (2b) we adopt the widespread idea (Rivero 1994a, b; Graffi 1996; Zanuttini 1997) that imperatival clauses fail to project any T-related functional structure, with the verb, together with any accompanying clitics, raising to the C-domain (cf. Rivero & Terzi 1995; Manzini & Savoia 2005 III: 388). In particular, we assume that in (2b) the verb and clitic *lu* first raise to a low C-related head (e.g. Fin°), from which the verb subsequently excorporates to reach a higher C-related head (e.g. Force°), presumably licensing imperatival illocutionary force. Consequently, the verb and the now enclitic object pronoun come to occupy adjacent but distinct head positions. Finally, in (2c) the complement of the coordinating conjunction is a bare nominal, in structural terms an NP embedded within a DP with a null head, such that between head and complement there intervenes a functional projection. In all three cases strict structural adjacency between Word$_1$ and Word$_2$ fails to obtain, but the configurations they engender are nonetheless sufficiently local to license RF in each case.

2.2.2   *Phases*

Evidence like that just considered in (2a-c) forces us to rethink the relevant locality conditions on the distribution of RF which appear to lend themselves more readily to an interpretation in terms of phase theory, rather than the more restrictive predictions of the three traditional core configurations in (1a-c). Indeed, examples like (2a-c) support the idea that the only relevant locality restriction on the adjacency requirement of RF is that Word$_1$ and Word$_2$ both surface in the same

phasal domain, a conclusion still in line with Fanciullo's original representational intuition about the 'minimal phrase'.

That this is indeed the correct conclusion is further demonstrated by examples (3a-c) which highlight how, although Word$_1$ and Word$_2$ are not immediately adjacent in structural terms, licensing of RF is again ensured by their co-occurrence within the same phase.

(3) a. (Un ssacciu) [$_{FocP}$ [$_{Spec}$ Cchi] [$_{Foc'}$ Foc° __ [$_{FinP}$ [$_{TP}$ ...[$_{AspP}$ **cc**unta
    not know.PRES.1SG   what   says
    Ccicciu(?)]]]]]](Spec…Head).
    Cicciu.
    'What is Cicciu saying? (I don't know what Cicciu is saying)'.

b. [$_{ModP}$ Pò…   [$_{AspP}$
        can.3SG   leave.INF **pp**arta [$_{v-VP}$ ~~parta~~]]] (Head…Head)
    'He can leave'.

c. [$_{ModP}$ Vò…   [$_{AspP}$ [$_{Spec}$
        want.PRES.3SG everything **tt**uttu]]] (Head…Comp)
    'He wants everything'.

For instance, (3a) illustrates a Spec…Head configuration where Word$_1$ lexicalizes a specifier position and Word$_2$ a lower head position, between which there intervene several projections. In particular, the Wh-phrase raises to SpecFocP, but the verbal head only raises to a head within Cinque's (1999) aspectual field above *v*-VP (Ledgeway & Lombardi 2005), witness, among other things, the ungrammaticality of subject-verb inversion in (direct) interrogatives which would be expected if the verb moved to the C-domain. Nonetheless, RF applies since Spec and Head are both contained within the higher CP-phase. Similarly, in the Head… Head configuration in (3b) Word$_1$ *pò* lexicalizes a modal head in the higher portion of Cinque's (1999) articulated clause structure, whereas the lexical infinitive instantiated by Word$_2$ raises to a clause-medial position (Ledgeway & Lombardi 2005: § 3.2) which, although at some structural distance from the former (cf. *Pò fforse/ssempe parta* 'He can perhaps/always leave'), still occurs in the same phase as it. Finally, (3c) illustrates a Head…Comp configuration in which the modal *vò*, once again merged in Cinque's higher modal field, finds itself at some considerable distance from its complement, the bare quantifier *tuttu* which lexicalizes the specifier of a plural completive aspectual projection situated in the lower portion of the aspectual field.

Furthermore, this latter example highlights how RF does not necessarily have to apply to the head of the complement (here not lexicalized), but simply applies to the closest adjacent item contained within the complement, its specifier in this case. This explains why in the following sequences we can find RF on the head of

the complement (4a), on its specifier (4b) or on a postnominal adjunct (4c). As a consequence, we can also find both Spec-Spec (5a) and Spec…Spec configurations (5b) where, as in the previous examples, Word₁ and Word₂ do not necessarily form a constituent (*pace* Fanciullo 1986), but RF signals their superficial linear adjacency within the same phase.

(4) a. [$_{NumP}$ tri   [$_{NP}$ **gg**atti]]
            'three   cats'

   b. [$_{NumP}$ tri   [$_{NP}$ [$_{Spec}$ **pp**overi] gatti]]
            'three   poor   cats'

   c. Ni   tiegnu   [$_{NumP}$ tri [$_{NP}$ [$_{N'}$ [~~ni~~] [$_{AP}$ **nn**ivuri]]]]
       thereof= have.PRES.1SG three   black.MPL
       'I've got three black ones'.

(5) a. [$_{DP}$ [$_{Spec}$ ogni] [$_{D'}$ D° — [$_{NP}$ [$_{Spec}$ **pp**overu] gattu]]]
            'every   poor   cat'

   b. [$_{TP}$ Un nni [$_{AspTermP}$ [$_{Spec}$ cchiù]   [$_{AspConP}$…[$_{AspPerfP}$ [$_{Spec}$ **ss**empe]
       not us=   anymore   always
       chiama]]]].
       call.PRES.3SG.
       'He no longer always call us'.

By the same token, we predict that RF should not obtain whenever Word₁ and Word₂ surface in distinct phasal domains, a prediction borne out by contrasts like (6a-b)-(7a-b).

(6) a. [$_{TopP}$ [$_{Spec}$ Accussì] [$_{Top'}$…[$_{TP}$ **pp**arava   ccu mia]]]
            thus   speak.PAST.3SG with me
       'He used to speak to me like that'.

   b. [$_{ForceP}$ [$_{Spec}$ Accussì] [$_{Force'}$…[$_{TP}$ parrava   ccu mia]]]
            thus   speak.PAST.3SG with me
       'Therefore he used to speak to me'

(7) a. [$_{ForceP}$ [$_{Force'}$ Fa-   [$_{FinP}$ [$_{Fin'}$ [$_{Fin°}$ [$_{Cl}$ llu]   [$_{Fin}$ ~~fa~~]
            do.IMP.2SG   =it.ACC.MSG
       [$_{v\text{-}VP}$ ~~lu fa~~!]]]]]] (Head–Head)
       'Do it!'

   b. [$_{ForceP}$ [$_{Force'}$ Fa   [$_{FinP}$ [$_{Fin'}$ ~~fa~~ [$_{TP}$ … tuttu   [$_{v\text{-}VP}$ ~~fa~~!]]]]]]]
            do.IMP.2SG   everything
       'Do everything!'

In (6a) *accussì* receives a topical reading, namely 'like that', and *qua* modal adverb lexicalizes SpecTop contained within the same (lower) CP phase as the finite verb *pparava* within the T-domain which bears RF. In (6b), by contrast, the verb no

longer bears RF and *accussì* is now a discourse connector interpreted as a resultative adverb 'therefore' external to the clause. Under this parenthetical use (cf. Cinque 1999: § 1.6), we assume that the adverb is merged in a high C-related position, in (6b) labelled as SpecForce for expository simplicity (but see Corr 2017 for detailed discussion of the syntax of the utterance beyond the canonical CP-layer), and, in particular, in a higher phasal CP domain from that in which the finite verb surfaces. As a consequence, RF fails to obtain since adverb and verb are contained within distinct phases and are sent to Spell-Out in distinct cycles of the derivation.

A similar contrast is seen in (7a-b). As already noted in the discussion of (2b) above, in (7a) the imperatival verb excorporates from Fin° to reach a higher C-related (e.g. Force°) head, a movement we assume to also characterize (7b). However, only in the former example does the imperatival verb license RF (for identical facts in the dialect of Campobasso, see D'Ovidio 1874: 179–180), a difference we can straightforwardly derive from the Phase Impenetrability Condition (PIC). In (7a) the imperatival verb and the enclitic *lu* surface in distinct phases instantiated by ForceP and FinP, respectively, but the verb continues to license RF on its pronominal complement since the latter occurs in the left edge (viz. head) of the FinP and hence remains accessible to the verb with which it is sent to PF in the same higher cycle.[5] In (7b), by contrast, the verbal complement *tuttu* occurs in the lower portion of the pre-VP space contained within the lower FinP phase, hence inaccessible to the potential RF effects of the verb in the higher ForceP.

Note finally that in (7b) the intermediate position (viz. copy) of *fa* in Fin° is not able to trigger RF on *tuttu*, despite both items occurring within the same phase at this point in the derivation. This is a general property of unpronounced copies which invariably fail to license RF, witness also example (8) where the intermediate copy of the RF trigger *cchi* 'what? in the embedded left periphery is unable to act upon the complementizer *ca*, although it occurs in a very narrow local Spec-Head configuration with the latter.

(8) Cchi bbu   [$_{CP}$ [$_{Spec}$ ~~cchi~~] [$_{C'}$ (*c)ca [$_{TP}$ facimu   ~~cchi~~?]]]
       what want.PRES.2SG   that   do.PRES.1PL
       'What do you want us to do?'

---

5. An anonymous reviewer points out that RF on the 3rd person accusative enclitic pronoun in (7a) might be explained by the fact that 3rd person accusative clitics always show a long lateral when enclitic to a (mono- or polysyllabic) imperative. This is, however, not empirically correct, inasmuch as 3rd person accusative enclitics show a non-geminate lateral following imperatives which do not trigger RF (e.g. *faciti(*l)lu!* 'do.IMP.2PL = it.ACC.MSG!'), and non-3rd person clitics also show lengthening after the 2SG imperative *fa* (e.g. *fammi/fatti santu!* 'make.IMP.2SG=me/you.SG saint!').

Indeed, example (8) contrasts with the indirect interrogative variant of (3a) above where we saw that the raised RF trigger *cchi* licensed RF on the embedded verb, despite not entering into a narrow Spec-Head relation with the latter which lexicalizes a much lower position within the T-domain yielding what we have informally been calling a Spec…Head configuration. As expected, examples like (7b) and (8) therefore confirm the view that (the features of) lower copies do not survive to PF, but, rather, represent simple bundles of formal features which, while playing a role in the conceptual-intentional system, do not play any role in the sensory-motor system to which they are invisible once all their phonological features have been stripped away.

3.   Summary and conclusions

The above discussion, although admittedly providing nothing more than a brief overview, has nonetheless highlighted how: (1) the phonological component does not necessarily operate in isolation, inasmuch as the distribution of Cosentino RF proves sensitive to syntactic structure offering us an important window on the interfaces, with PF directly externalizing not only syntactic but also pragmatico-semantic information such as that observed in the contrast between (6a-b) (cf. Manzini & Savoia 2016: 413)[6]; (2) the syntactic constraints operative on the distribution of RF should not be interpreted in terms of three distinct and unrelated structural configurations (viz. Spec-Head, Head–Head and Head-Comp) but, can be more accurately, simply and elegantly understood and modelled in terms of a unitary phase-based approach; (3) not only does phase theory offer us a more adequate understanding of the structural restrictions on the application of Cosentino RF, but the distribution of the latter also provides us with new and interesting data to test the nature and computation of phasal domains, witness, for example, our conclusion that the postulation of a single CP phase is not sufficient to account for examples such as (6)–(7) which force us to recognise distinct lower and higher phasal domains within the C-domain.

In short, it is clear that the distribution of RF in Cosentino, as well as in other dialects of southern Italy which show considerable subtle but structured

---

6. For a discussion of the role of Cosentino RF in signalling informational structure content and coreference relations, see Ledgeway (in prep.). Note, however, that syntactic information is not necessarily always externalized at PF, as highlighted by D'Alessandro & Scheer (2015) who in their modular approach to phase theory, in which Spell-Out is separated from the PIC, provide convincing evidence for the claim that the PIC does not necessarily apply uniformly in both syntax and phonology.

microvariation in the licensing and surface effects of RF, offer many important opportunities to deepen our understanding of the phonological correlates of syntactic (and pragmatico-semantic) representations. Consider, for instance, the structurally-determined alternations in the distribution of RF following the singular persons of the perfective auxiliary and copular uses of BE (viz. 1/2/3SG SUM/SIS/EST > *so/si/è*) widespread in southern dialects as witnessed in the numerous paradigms reported in Manzini & Savoia (2005, II-III: Chapters 5–6; cf. also Torcolacci 2014a, b), a small selection of which are summarized in Table 1.

Table 1. Distribution of RF following singular persons of BE

	Transitive			Unaccusative			Reflexive			Copula		
	1	2	3	1	2	3	1	2	3	1	2	3
Poggio Imperiale	+	+	+	+	+	+	+	+	+		+	+
S.Benedetto del Tronto	−	−		−	−		−	−	−	+	+	+
Pontecorvo	−	+	−	−	+	−	−	+	−	+	+	+
Pàstena-Castelpetroso	−	−	−	+	+	+	+	+	+	+	−	+
Ruvo di Puglia	−	−	−	+	−	+	+	−	+	+	−	+
Tufillo	−	−		+	+	+	−	−	−	?	?	?
Secinaro	−	+	+	−	+	+	−	+	+	+	+	+

Table 1 distinguishes between auxiliary and copula uses of BE in which it is respectively followed by a transitive, unaccusative and reflexive participle and by an adjective. ± indicate the presence/absence of RF, ? the absence of relevant information in Manzini & Savoia (2005), and an absence of any symbol indicates that the relevant potential RF trigger is not licensed in that particular person/context.

The data reported in Table 1 are far too complex for us to do justice to them here, requiring further detailed investigation of the individual microparametric properties of each of the varieties concerned. Suffice it to note here some significant patterns which deserve future investigation. Consider, for instance, the northern Pugliese dialect of Poggio Imperiale (Manzini & Savoia 2005: 720–721) where the forms of BE for all three persons are RF triggers and systematically license RF, whenever selected, in all four contexts. One possible way to interpret this distribution is to assume, for instance, that in Poggioimperialese finite V-movement is invariably very low such that all forms of BE, whether auxiliary or copula, remain within *v*-VP and hence within the same phase as their participial or adjectival complement. Quite different is the behaviour of the southern Marchigiano dialect of San Benedetto del Tronto (Manzini & Savoia 2005: 682–683) where we witness an active-stative split, in that forms of BE never trigger RF on a following active participle, but

consistently license RF in conjunction with an adjectival complement.[7] On the standard assumption that active *v*Ps are phasal but the stative *v*P instantiated by copular BE is not, the observed RF contrast follows straightforwardly. If we further allow that the phasal status of different *v*P types is subject to cross-linguistic parametric variation – also in terms of the variable application of the PIC in syntax and phonology, as argued in D'Alessandro & Scheer (2015) –, then we can also account for the contrast in dialects like Pàstena-Castelpetroso (Manzini & Savoia 2005: 713–714) where a transitive-unaccusative split obtains with RF consistently blocked solely before transitive participles.[8] Once again, on the uncontroversial assumption that θ-complete transitive *v* is a (strong) phase head, whereas non-θ-complete unaccusative and copular *v*, with which reflexive predicates are aligned in this variety, is not, then the observed distributional contrast in RF follows without further stipulation. However, we know from Romance auxiliary selection (cf. Ledgeway in press: § 4.3), among other things, that (different classes of) reflexive predicates can variously align with both transitives or unaccusatives in different Romance varieties. Thus, we expect to also find varieties such as the southern Abruzzese dialect of Tufillo (Manzini & Savoia 2005: 747–748) where RF is blocked not just before transitive but also before reflexive participles, but is systematically licensed in conjunction with unaccusative participles. This type of approach further allows us to understand otherwise puzzling contrasts such as those in (9a-c) from the northern Abruzzese dialect of Secinaro (Manzini & Savoia 2005, II-III: 691, 113).

(9) a. sɔ   kɹunˈtintə (Secinaro)
         be.PRES.1SG happy
         'I am happy'.

    b. (mə) sɔ   parˈlaːtə / məˈnʊːtə / laˈvaːtə (Secinaro)
        me= be.PRES.1SG spoken   come   washed
        'I have spoken / come / washed (myself)'.

    c. sɔ   tːə   camˈatə / sɔ   mːə laˈvatə (Secinaro)
        be.PRES.1SG =you.SG called   be.PRES.1SG =me washed
        'I have called you / I have washed (myself)'.

The distribution of RF following 1SG sɔ in Secinaro reveals an active-stative split, with RF systematically licensed in conjunction with an adjectival complement (9a), but otherwise blocked before an active participle (9b). Assuming the active

---

7. Cf. a similar active-stative distribution in the southern Laziale dialect of Pontecorvo (Manzini & Savoia 2005: 701–702), albeit limited to the 1SG and 3SG.

8. This transitive-unaccusative split is paralleled by the Pugliese dialect of Ruvo di Puglia (Manzini & Savoia 2005: 724), albeit limited to the 1SG and 3SG.

participles in (9b) to lexicalize a low position within the *v*-VP complex, then the failure of 1SG sɔ to trigger RF on the participle follows from the phasal status of the *v*P containing the latter. In (9c), by contrast, the 1SG RF trigger is now followed by a clitic pronoun which we standardly take to have vacated the *v*-VP complex to reach a T-related position, from which auxiliary sɔ subsequently excorporates yielding enclisis (cf. analysis of (2b) above). In contrast to (9b), 1SG sɔ now licenses RF in (9c) since it and the clitic both surface in the same higher CP phase, albeit in distinct functional heads.

Clearly, the finer details of the analyses outlined here based on differing parametric instantiations of phasal domains remain to be worked out, but they certainly provide a possible way forward to understanding the structural regularities that underlie what might otherwise be written off as superficial phonological irregularities.

Acknowlegments.

References

Andalò, Adrina. 1991. Il raddoppiamento sintattico nel dialetto di Napoli. In *Tra rinascimento e strutture attuali. Saggi di linguistica italiana*, Luciano Giannelli et al. (eds) 241–251. Turin: Rosenberg & Sellier.

Biberauer, Theresa & D'Alessandro, Roberta. 2006. Syntactic doubling and the encoding of voice in eastern Abruzzese. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, Donald Baumer, David Montero & Michael Scanlon (eds) 87–95. Somverille MA: Cascadilla Proceedings Project.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A Cross-linguistic Perspective*. Oxford: OUP.

Corr, Alice. 2017. Ibero-Romance and the Syntax of the Utterance. PhD dissertation, University of Cambridge.

D'Alessandro, Roberta & Scheer, Tobias. 2015. Modular PIC, *Linguistic Inquiry* 46(4): 593–624. https://doi.org/10.1162/LING_a_00195

D'Ovidio, Francesco. 1874. Fonetica del dialetto di Campobasso. *Archivio Glottologico Italiano* 4(1): 145–184.

Fanciullo, Franco. 1986. Syntactic reduplication and the Italian dialects of the centre-south. *Journal of Italian Linguistics* 8: 67–103.

Fanciullo, Franco. 1997. *Raddoppiamento sintattico e ricostruzione linguistica nel sud italiano*. Pisa: ETS.

Graffi, Giorgio. 1996. Alcune riflessioni sugli imperativi italiani. In *Italiano e dialetto nel tempo. Saggi di grammatica per Giulio C. Lepschy*, Paola Benincà, Guglielmo Cinque, Tullio De Mauro & Nigel Vincent (eds) 133–148. Rome: Bulzoni.

Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*.[Beihefte zur Zeitschrift für romanische Philologie Band 350] Tübingen: Max Niemeyer.
https://doi.org/10.1515/9783484971288

Ledgeway, Adam. 2016a. Italian, Tuscan, and Corsican. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 206–227. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199677108.003.0014

Ledgeway, Adam. 2016b. The dialects of southern Italy. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 246–269. Oxford: OUP.
https://doi.org/10.1093/acprof:oso/9780199677108.003.0016

Ledgeway, Adam. In press. Parameters in the development of Romance perfective auxiliary selection. In *Historical Linguistics 2015*. Selected Papers from the 22nd International Conference on Historical Linguistics, Naples, 27–31 July 2015, Michela Cennamo & Claudia Fabrizio (eds) Amsterdam: John Benjamins.

Ledgeway, Adam. In prep. The syntactic distribution of *raddoppiamento fonosinttatico* in Cosentino: A phase-theoretic account. *Romance Languages and Linguistics Theory*, Dragomirescu, Adina & Nicolae, Alexandru (eds), Amsterdam: Benjamins.

Ledgeway, Adam & Lombardi, Alessandra. 2005. Verb movement, adverbs and clitic positions in Romance. *Probus* 17: 79–113. https://doi.org/10.1515/prbs.2005.17.1.79

Loporcaro, Michele. 1988. History and geography of *raddoppiamento fonosintattico*: Remarks on the evolution of a phonological rule. In *Certamen Phonologicum*, Pier Marco Bertinetto & Michele Loporcaro (eds), 341–387. Turin: Rosenberg & Sellier.

Loporcaro, Michele. 1997. Lengthening and *raddoppiamento fonosintattico*. In *The Dialects of Italy*, Martin Maiden & Mair Parry (eds) 41–51. London: Routledge.

Manzini, M. Rita & Savoia, Leonardo. (2005). *I dialetti italiani e romanci. Morfosintassi generativa.*, 3 Vols. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Savoia, Leonardo. 2016. Quale informazione sintattica viene interpretata dai fenomeni fonosintattici? Evidenza dalle varietà italiane meridionali. In *Actes du XXVIIe Congrès international de linguistique et de philologie romanes, (Nancy, 15–20 juillet 2013), Section 4: Syntaxe*, Adam Ledgeway, Michela Cennamo & Guido Mensching (eds) 403–414. Nancy: ATILF. <http://www.atilf.fr/cilpr2013/actes/section-4.html>

Rivero, María-Luisa. 1994a. Negation, imperatives and Wackernagel effects. *Rivista di Linguistica* 6: 39–66.

Rivero, María-Luisa. 1994b. Clause structure and V movement in the languages of the Balkans. *Natural Language and Linguistic Theory* 12: 63–120. https://doi.org/10.1007/BF00992746

Rivero, María-Luisa & Terzi, Arhonto. 1995. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31: 301–332. https://doi.org/10.1017/S0022226700015620

Rizzi, Luigi & Savoia, Leonardo. 1993. Conditions on /u/ propagation in southern Italian dialects. In *Syntactic Theory and the Dialects of Italy*, Adriana Belletti (ed.) 252–318. Turin: Rosenberg & Sellier.

Rohlfs, Gerhard. 1966. *Grammatica della lingua italiana e dei suoi dialetti. Fonetica*,Vol. 1. Turin: Einaudi.

Rohlfs, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti. Sintassi e formazione delle parole*, Vol. 3. Turin: Einaudi.

Sampson, Rodney. 2016. Sandhi phenomena. In *The Oxford Guide to the Romance Languages*, Adam Ledgeway & Martin Maiden (eds) 669–680. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677108.003.0040

Savoia, Leonardo M. 1987. Teoria generativa, modelli fonologici e dialettologia. La propagazione di u in una varietà lucana. *Rivista Italiana di Dialettologia* XI: 185–263.

Savoia, Leonardo M. 2015. *I dialetti italiani. Sistemi e processi fonologici nelle varietà di area italiana e romancia*. Pisa: Pacini.

Silvestri, Giuseppina. 2007. Il rafforzamento fonosintattico nel dialetto di Verbicaro (Cosenza). *L'Italia Dialettale* 68: 146–151.

Silvestri, Giuseppina. 2014. *Rafforzamento fonosintattico as a phenomenon of the phonology-syntax interface*. Paper presented at the *Romance Linguistics Seminar*, University of Cambridge, 28 October.

Torcolacci, Giuseppe. 2014a. Il raddoppiamento fonosintattico e la codifica di tratti morfosintattici. Il caso dei dialetti italiani meridionali.,In *L'Italia Dialettale. In Approcci diversi alla dialettologia italiana contemporanea*. Roberta D'Alessandro, Claudio Di Felice, Irene Franco & Adam Ledgeway (eds). Special issue of*L'Italia Dialettale* 75: 247–271.

Torcolacci, Giuseppe. 2014b. *Marking the Default. Auxiliary Selection in Southern Italian Dialects*. Leiden: LOT.

Vincent, Nigel. 1988. Non-linear phonology in diachronic perspective: Stress and word-structure in Latin and Italian. In Certamen Phonologicum, Pier Marco Bertinetto & Michele Loporcaro (eds), 421–432. Turin: Rosenberg & Sellier.

Zanuttini, Raffaella. 1997. *Syntactic Properties of Sentential Negation. A Comparative Study of Romance Languages*. Oxford: OUP.

Metaphony as magnetism

Marc van Oostendorp and Roberta D'Alessandro

Meertens Instituut and Radboud University / Utrecht University

Metaphony in Romance poses a well-known problem for Element Theory, as it seems to involve lowering. [D'Alessandro and van Oostendorp (2016)] propose to solve this by assuming some suffixes are ' |A| Eaters', absorbing the |A| element from the stem vowel without getting phonetically realized themselves. This paper points to some problems with this analysis, and shows that Magnetic Grammar, a framework in which all linguistic variation is encoded in features, might help to solve them.

Keywords: Element Theory, phonology, primitives, Magnetic Grammar

1.   Introduction

The literature on the phonology of many languages knows a few classics: topics to which every student of the language will one day turn. For the languages of Italy, metaphony is clearly such a topic [Savoia (2005), Savoia (2015), Calabrese (2011), Torres-Tamarit et al.(2016)]. Metaphony is a phenomenon that is found in a rather large number of varieties (also in some Romance varieties of the Iberian peninsula, as well as, depending on one's definitions, in southern French). In those dialects, it shows a bewildering amount of variation; it is hardly an exaggeration to say that no two dialects have the same system of metaphony.

Furthermore, the process works at the interface of several modules: it is partly morphological, or even lexical, in at least some dialects, but it clearly works with phonological objects such as vocalic height, as descriptively metaphony involves vowel raising. The precise way in which raising works out can also differ from one dialect to the next. For instance, in some dialects low mid vowels /ɛ/ raise to [e], while in others they raise to [i]. And finally, subsegmental phonology also interacts with prosody in an interesting way: it is the stressed vowel of the stem that typically undergoes metaphony, sometimes skipping unstressed vowels in between the apparent trigger (the suffix vowel) and its target [Walker (2014), Mascaró, 2015)].

We do not think that anybody can claim to have a complete theory covering all cases of metaphony, if only because not all cases have been adequately described yet. In this short squib, we do not aim to solve the puzzle either, but wish to show how it is relevant for some current thinking about the structure and status of the primitives of our representations: features or elements. We argue that metaphony does not just act as a magnet on researchers, it is indeed itself a result of grammatical magnetism.

2.   Metaphony and the |A| eater

[D'Alessandro and van Oostendorp (2016)] discuss a puzzle for standard Element Theory [Backley (2011)]: how to represent metaphony in Italo-Romance [Savoia and Maiden (1997)]. The general problem here is that metaphony involves a morphological process that raises (stressed) vowels in certain morphological contexts, e.g. in the plural of nouns. E.g. in Grado [Walker (2005)], we find alternations such as the following:

(1) a. *tempo* 'time', *timpi* 'times'
     b. *fjor* 'flower', *fjuri* 'flowers'

Plural formation thus consists of two parts: a suffix is added which in this case is *-i*, and metaphony is applied to the root vowel, so that [e] turns into [i] and [o] into [u]. The problem is that such changes cannot easily be described in terms of Element Theory, which has no primitive corresponding to 'high vowel'. The full representation of [i] is |I|, and the full representation of [u] is |U| and these two have nothing in common, so there is nothing which could, for instance, 'spread' from the suffix to the stem vowel to get the required result.

The problem is particularly salient in dialects such as that spoken in Arielli (Abruzzese) in which metaphony is triggered in forms which have no clear suffix marking at all. The following are all masculine nouns – metaphony does not apply in the feminine:

(2) a > i lu canə   li chinə   'the dog/s'
     e > i lu velə   li vilə   'the veil/s'
     ɛ > i lu martɛllə li martillə   the hammer/s
     ɔ > u lu nəpɔtə   li nəputə   'the nephew'
     o > u lu waʎ:onə   li waʎ:unə   'the young boys'

All round vowels turn to [u] in the plural, all other vowels turn to [i]. We find similar alternations also in adjectives, as well as in verbs (e.g. in the second person singular).

Within Element Theory, one needs to say that metaphony somehow subtracts the |A| element. Arielli shows this cannot plausibly be the result of any overt phonological material present in the suffixes involved: as a result of historical reduction, the singular and plural forms all end in a schwa. Metaphony thus is the only exponent of plurality for these masculine nouns.

[D'Alessandro and van Oostendorp (2016)] present an analysis of this fact in terms of what they call an ' |A| Eater': an empty mora which is an exponent of the plural and which is suffixed to the word. The A element of the root is then shifted to this mora, but the mora itself is not linked to any higher-order structure. Because only material that is linked to the phonological word is pronounced, the mora is not and neither is the |A| element linked under it. We use σ as the shorthand for whatever the syllabic constituency is:

(3)
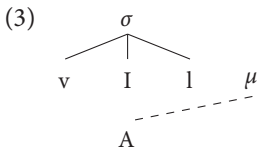
This analysis sees metaphony thus as a kind of truncation (of the |A| element), and analyses this in line with [Trommer and Zimmermann (2014), Zimmermann (2017)]'s view on this phenomenon in such a way that it fits an Items-and-Arrangement view of morphology [Hockett (1958)]: a morphologically more complex form (such as the plural) always has more phonological structure than a simpler form (such as the singular). It may look as if the plural is really missing something (the |A| element), but as a matter of fact it has something extra (the mora that is eating the |A| element).

[D'Alessandro and van Oostendorp (2016)] point to an interesting piece of evidence for the |A| Eater: it shows up in certain phrasal contexts, for instance before postnominal adjectives:

(4)   a.   *lu canə* 'the dog'
      b.   *li chinə* 'the dogs'
      c.   *lu canə cioppə* 'the lame dog' (* *lu cana cioppə*)
      d.   *li china ciuppə* 'the lame dogs'

(4d) has an [a] that is not etymological. The question is where it is coming from, and [D'Alessandro and van Oostendorp (2016)] argue that it is an overt manifestation of the |A| Eater, which stays uninterpreted at edges of constituents, but can make it to the surface when it is followed by other material.

[D'Alessandro and van Oostendorp (2016)] thus give a nice analysis of Arielli metaphony, but there are some problems with the precise representation of the

|A| Eater. If this is an empty mora, the question is whether all empty moras will always show |A| -eating behaviour. An implicit answer in [D'Alessandro and van Oostendorp (2016)] is that this is under grammatical control (the analysis is more or less set in an OT frame where some constraint could be made responsible for this), but that is obviously not directly in line with the assumptions of Government Theory which would say that similar representations should trigger similar behaviour in different languages.

This first problem may still be solved, but more serious is the question how we would represent metaphony in those Italian dialects in which the ending triggering metaphony is not reduced, but rather a full vowel, such as in the Grado dialect. [D'Alessandro and van Oostendorp (2016)] suggest that the emptiness of the suffix in Arielli is a key factor in its attracting an element from the root: empty vowels are not allowed. The ending here cannot be an empty mora, since it is a full vowel [i]. Still, this suffix must somehow be able to attract the |A| element.

Finally, there is another problem left explicitly unsolved in [D'Alessandro and van Oostendorp (2016)]: why the plural of *cane* is *chine* with an [i]. The latter vowel apparently comes out of nowhere. [D'Alessandro and van Oostendorp (2016)] note that it may be the default epenthetic vowel: since stressed schwa is not allowed, the language fills in an |I| element by default. It is however hard to come by evidence that the |I| has this epenthetic function also elsewhere in the language.

3. Magnetic grammar

[D'Alessandro and van Oostendorp (2017)] argue for a view of variation in grammar which is based purely on features and elements, which they call Magnetic Grammar (MG). The assumption is that all variation between languages is in the representation of the primitives. These are assumed to be (monovalent) features, but the actual content of the primitives is irrelevant, so that the idea would work as well for elements.

The idea is that languages, first, obviously choose a set of such primitives from a universal set. A language that has no voiced consonants and no low tones does not use the |L| element; this is something which the language-learning child has to learn – learning that one's language does not have contrastive voicing or contrastive tone *equals* learning that there is never an |L| element in any representation. This assumption is probably shared by all versions of feature and element theory.

Also the assumption that features can be combined in a language-specific way is not unique to MG. A language that does not allow for front rounded vowels, does not allow a combination of |I| and |U|; this is usually called a (parametrized) 'licensing constraint' in the literature on Element Theory [Charette and Göksel

(1996)]. Magnetic Grammar proposes, however, that the combinability of an element with another element is not somewhere in the grammar, but rather a language-specific property of the elements itself. If a language does not have front rounded vowels, its |I| has a diacritic indicating that it cannot be combined with U: $|I|_{*U}$ (the language may also have |U| diacriticed as $|U|_{*I}$; the difference between those options will follow from other considerations about the language).

We read the diacritic X_{*Y} as 'X repels Y'. Next to repulsion, we assume there can also be a force of attraction between two elements. For any element X in a language that also has an element Y, there are thus three possibilities:

(5) a. X_{*Y}: X repels Y, a representation is ill-formed if X and Y occur in the same domain

b. $X_{\supset Y}$: X attracts Y, a representation is ill-formed if X and Y do not occur in the same domain

c. X: there is no relation from X and Y, a representation which is well-formed regardless of whether X and Y occur in the same domain (modulo possible attracting or repelling properties of Y or other elements)

MG has it that phonological (and syntactic) representations consist of such primitives, which are combined in higher-order consitituents (segments, syllabic constituency in the case of phonology, phrase structure in the case of syntax); the (only) difference between languages is the set of primitives, including their diacritics on combinability.

In line with the work of (among others) [Charette and Goksel(1996)], MG holds that these properties do not just govern static inventories of segments, but also 'active' phonological processes. For instance, Turkish has an eight vowel system, which can of course be easily described if we assume that the three vowel primitives can be freely combined (and all of them are combinable with the basic vocalic primitive, which we will write here as |@|.).

(6) ɪ i u y
     @ @ @ @
        I   I
           U U
     a e o ø
     @ @ @ @
        I   I
           U U
     A A A A

This gives us initially, the following view on the Element inventory:

(7)   @, I, U, A

As is well-known, Turkish has front harmony, meaning that suffix vowels (and other unspecified vowels, such as those that are the result of epenthesis) will be front after a stem ending in a front vowel. (Vowel harmony is of course much more complicated: it also involves roundedness harmony and all kinds of restrictions of application, which we will not discuss here, as we introduce the phenomenon here merely for illustrative purposes; see [Kabak (2011)]).

We can solve this by assuming that the |@| element attracts frontness (at least when it occurs in a suffix):

(8) $@_{\supset I}$

This means that the representations for {ı, a, u, o} are all insufficient as they are lacking an |I| element. Grammar provides two ways to solve this: by spreading an |I| from a neighbouring segment, or by inserting a new |I| element. If the stem does not contain a front vowel, the former option is not possible, and hence the only possibility would be to insert an |I| element.

Yet this, in turn, is apparently not allowed in Turkish: a back vowel only turns front if there is another front vowel in the word [the following examples are from Kabak:2011]

(9)   a.   *janluʃ* 'wrong' (* janliʃ, * jenluʃ)
      b.   *josun* 'moss' (*josyn, * jøsun)
      c.   *zengin* 'rich' (* zengyn, *zangin)
      d.   *kømyr* 'coal' (*komyr, *kømur)

This means that 'epenthetic' elements are not allowed. That may be a language-specific property, in which case we have to represent it in some way. One way to do it is to say that an element has to be linked to a lexical category:

(10) $I_{\supset X}$, where X is a categorial feature such as N, V.

If a vowel belongs to the phonological exponence of a noun, all of its elements will be inherently linked to the categorial (and other morphosyntactic) features of that noun, and therefore satisfy this condition. However an epenthetic vowel has no connection to any morphosyntactic feature, and therefore will not satisfy this condition. In this way, we thus implement the ideas of Coloured Containment [van Oostendorp (2007)] in Magnetic Grammar.

We thus can revise (7) to the following:

(11) $@_{\supset P}$, $I_{\supset N}$, U, A

It should be noted that the requirements on |@| and |I| are in conflict if the nominal stem does not contain an underlying |I| element. In that case, apparently in Turkish the requirement on |I| 'wins', in the sense that we end up having non-fronted vowels, so that @$_{\supset I}$ stays unsatisfied. Our suggestion is that only the 'weak' 'default' element |@| (the 'cold vowel') can have such a weak requirement that is not fulfilled if it conflicts with the requirements of other elements. We thus do not need the kind of language-specific ranking of constraints that is so well-known from Optimality Theory, although we have something that is minimally equivalent: first, there is only a difference between 'weak' elements and 'strong elements', not a complete ranking of all elements; and secondly we expect properties of e.g. |@| always to be 'weak' as this element is also weak in other ways (it does not influence the phonetic profile of the vowel involved, for instance).

4.   Metaphony in magnetic grammar

Since metaphony is sometimes compared to vowel harmony, one could expect it to be represented in a similar way. Of course the restrictions of Element Theory still apply: we cannot say that the stressed stem vowels undergo metaphony because they attract some feature from the suffix, since the suffix does not have such a feature (at least in Arielli, where it is schwa = |@|, which is already contained in the other vowels by definition), and because the result of metaphony is still *subtraction* of the |A| element. It also cannot be a floating element, as there simply is not element which corresponds to the valued feature [−high] .

   The MC framework provides us with a way to improve on [D'Alessandro and van Oostendorp (2016)]'s implementation of the |A| Eater. We no longer need to say that a suffix has to be empty to 'eat' an A element; it needs to be merely attracting it. In other words, the suffix can consist of a vowel that attracts an A. This could be the schwa element:

(12)   @$_{\supset A}$

We then get an analysis that is exactly parallel to the one in [D'Alessandro and van Oostendorp (2016)], but with the advantage that we have given a clearer representational reason why the plural affix can behave in this way, without having to introduce a constraint ranking or some other grammatical device that is specific to this morpheme. Like in the previous account, we still need to understand why the suffix [a] which is formed in this way only shows up if it is not peripheral, but we can account for this by relying on general phonotactic considerations of the language: also etymological [a] only shows up under those conditions [Passino(2016)].

This is of course the same kind of representation we have given to harmony, albeit that in this case it is the apparent *trigger* of the harmony (the suffix) that has the specification to attract the |A| element. This is again the weak cold vowel |@|, so that if there is no |A| present in the environment, the attraction property stays unsatisfied.

Notice that this schwa element can in principle be part of a larger structure. In particular, it can be combined with an |I| element. This means that in isolation, it would be pronounced as an [i] vowel, just like it is pronounced in a northern dialect like Grado. Given the fact that southern dialects like Arielli have a general reduction of final vowels, probably also as an active synchronic process, the |I| element will not be pronounced in this dialect, so that we as a matter of fact could posit a uniform representation for the exponence of plural suffix in both (groups of) dialects:

(13)   + PLURAL $\Leftrightarrow$ | @$_{\supset A}$ I | (to be revised in (14))

However, this representation still faces the problem that after metaphony, the vowel never shows up as a combination of |I,A|, i.e. as *e*. In Grado, only the |I| shows up (*fjuri*), whereas in Arielli only the |A| surfaces, under the right conditions (*china ciupp*).

We propose that the reason is that the |@| element is not just attracting |A|, but also repelling |I|, in both dialects.

(14)   + PLURAL $\Leftrightarrow$ | @$_{\supset A, *I}$ I |

Since the |I| cannot be realized in the same position as the schwa, it needs to find a different position. Given the workings of grammar, this position needs to be as close as possible. Grado chooses a position immediately adjacent to the suffix. Because the suffix is itself not realized, it looks as if [i] itself is the suffix:

(15)

$$\sigma$$

f   j   U   r   I   @$_{\supset A, *I}$

A

The reason why the I is pronounced rather than the schwa (and the A attached to it) may be, again, that schwa is too weak. In Arielli, this solution is not possible, as the final position only allows schwa. For this reason, it is the |I| which is left floating, giving a picture that is similar to (3):

(16)

Now the fact that it is an [i] that surfaces in the plural of *cane* (*chine*) falls into place. Since stressed schwas are not allowed in Arielli, the position needs to be filled by the only element that is free:

(17)

The |I| only moves to the stressed position when it is not already filled, possibly because the accent is marked in Arielli (as in most languages of Italy) as attracting elements that are not the cold vowel.

The |A| Eater now has become a rather terroristic morpheme: it grabs an |A| from the context when it can, but it also tries to get rid of its |I| element, either making it masquerade as a suffix (in Grado) or intrude into the root (in Arielli). It is this disruptive behaviour, we propose, which makes it into such an interesting topic of study.

5.   Conclusion

The idea that grammar is organized by forces of magnetic attraction and repulsion is currently a programme rather than a fully worked-out proposal. In this short contribution we have studied one of its implications: it allows us to give representations of inflectional endings exactly the force they seem to need in order to account for the behaviour of metaphony, including even the promise that different dialects might have a very similar representation of metaphony-triggering suffixes after all.

This then allows us to give a phonological analysis of metaphony. Morphology does nothing special, except filling in the slot for (e.g.) plurality with a phonological representation which then starts interacting with its phonological context. It is not necessary to take resort to extra-phonological devices and still account for the apparently rather strange behaviour of 'raising' suffixes which furthermore are not high themselves.

References

Backley, Phillip. 2011. *An Introduction to Element Theory*. Edinburgh: EUP.

Calabrese, Andrea. 2011. Metaphony in Romance. In *The Blackwell Companion to Phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume& Keren Rice (eds) 2631–2643. Malden MA: Blackwell-Wiley. https://doi.org/10.1002/9781444335262.wbctp0110

Charette, Monik& Göksel, Asli. 1996. Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 6: 1–25.

D'Alessandro, Roberta& van Oostendorp, Marc. 2016. Abruzzese metaphony and the a eater. In Torres-Tamarit, F., Linke, K., & van Oostendorp, M. (eds), 349–368.

D'Alessandro, Roberta& van Oostendorp, Marc. 2017. Magnetic grammar. *Why phonology isn't different, and neither is syntax*. Ms, Utrecht University and Meertens Instituut.

Hockett, Charles. 1958. Two models of grammatical description. *Word* 10: 210–231. https://doi.org/10.1080/00437956.1954.11659524

Kabak, Baris, 2011. Turkish vowel harmony. In *Blackwell Companion to Phonology*, Marc Van Oostendorp, Colin J. Ewen, Elizabeth Hume, & Keren Rice (eds) 2831–2854. Malden MA: Wiley-Blackwell. https://doi.org/10.1002/9781444335262.wbctp0118

Mascaró, Joan. 2015. Myopic harmony and nonlocal trigger-target asymmetries. Ms, Universitat Autonoma de Barcelona. <https://mascaro-uab.weebly.com/uploads/4/0/9/2/40925737/myopicharmony.pdf> (6 February 2018).

van Oostendorp, Marc. 2007. Derived environment effects and consistency of exponence. In *Freedom of Analysis?*, Sylvia Blaho, Patrik Bye& Martin Krämer (eds) 123–148. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110198591.123

Passino, Diana. 2016. Progressive metaphony in the Abruzzese dialect of Teramo. In *Approaches to Metaphony in the Languages of Italy*, Francesc Torres-Tamarit, Kathrin Linke& Marc van Oostendorp (eds) 123–145. Berlin: Mouton de Gruyter.

Savoia, Leonardo M.& Maiden, Martin. 1997. Metaphony. In *The Dialects of Italy*, Martin Maiden& Mair Parry (eds) 15–25. London: Routledge.

Savoia, Leonardo M. 2005. Armonizzazioni vocaliche in alcune varietà romanze. In *Italia linguistica: Discorsi di scritto e di parlato. Nuovi studi di linguistica italiana in onore di Giovanni Nencioni*, Marco Biffi, Omar Calabrese& Luciana Salibra (eds) 217–233. Siena: Protagon.

Savoia, Leonardo M., 2015. I dialetti italiani. Sistemi e processi fonologici nelle varietà italiane e romance. Pisa: Pacini.

Torres-Tamarit, Francesc, Linke, Kathrin& Marc van Oostendorp (eds). 2016. *Approaches to Metaphony in the Languages of Italy*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110366310

Trommer, Jochen& Zimmermann, Eva. 2014. Generalized mora affixation and quantity-manipulating morphology. *Phonology* 31(3): 463–510. https://doi.org/10.1017/S0952675714000220

Walker, Rachel. 2005. Weak triggers in vowel harmony. *Natural Language and Linguistic Theory* 23: 917–989. https://doi.org/10.1007/s11049-004-4562-z

Walker, Rachel. 2014. Nonlocal trigger-target relations. *Linguistic Inquiry* 45: 501–523. https://doi.org/10.1162/LING_a_00165

Zimmermann, Eva. 2017. *Morphological Length and Prosodically Defective Morphemes*. Oxford: OUP. https://doi.org/10.1093/oso/9780198747321.001.0001

Some reflections on the syllabification of clusters

A view from the dialects of Italy

Diana Passino
University of Côte d'Azur & CNRS, BCL

Throughout different theoretical models it is generally agreed that syllable structure is only determined by the melodic composition of segments and predictions are made concerning the syllabification of clusters on the basis of their segmental content. This contribution draws attention on empirical data showing that the predictions are not always verified, arguing in favour of a refinement of the theoretical tools proposed so far to derive cluster syllabification.

Keywords: syllable theory, dialects of Italy, consonant clusters

1.   Introduction

This contribution collects some reflections on the syllabification of consonant clusters focusing on how empirical data on syllabification match the predictions of standard syllable theory. Based on these reflections, I wish to argue that empirical data call for a revision of a number of standard assumptions concerning syllabification and a rethinking of some of the theoretical tools. To illustrate my argument I draw on dialects of Italy and dedicate these reflections to Leonardo Savoia and to his pioneering and far reaching work in this domain, whereby a truly extensive and thorough fieldwork is complemented with an acute formal analysis.

The contribution is organised as follows: Section 2 discusses issues concerning standard syllable theory and its predictions concerning the syllabification of clusters. Section 3 introduces and examines actual data relevant for the topic, mostly coming from the dialects of Italy. Section 4 offers some preliminary conclusions.

2. Syllable theory and the syllabification of clusters

There is general agreement across standard theories of the syllable in conceiving it as a unit that organizes speech sounds in terms of their intrinsic sonority, whereby each syllable corresponds to a sonority peak around which sounds are arranged. Sonority is roughly defined as the acoustic intensity or the loudness of a sound relative to the energy used to produce that sound (Blevins 1995: 211; Parker 2002). The tendency for segments to regroup monotonously rising in sonority from the onset of the syllable towards the peak and monotonously falling towards the end of the syllable (coda) is referred to as the Sonority Sequencing Principle or SSP (Selkirk 1984) and the relative sonority of segments as the Sonority Hierarchy. These modern theoretical tools incorporate findings that date back to Jespersen (1904), who established a link between syllable well-formedness and sonority-related ordering of sounds, as well as Sievers (1881), who proposed an arrangement of sounds in scales related to their intrinsic strength. In modern theorizing the SSP and its corollary Minimal Sonority Distance Principle (Steriade 1982), as well as the Maximal Onset Principle (henceforth MOP) and the principle according to which word-edges coincide with syllable-edges, are particularly helpful to obtain syllabification by means of an algorithm. The most sonorous segments, vowels, and in some languages also sonorants, are assigned to nuclei. As for the consonants, it is generally agreed that, by the application of the MOP, when a consonant immediately precedes a vowel it belongs to the same syllable as the vowel. A somewhat more difficult task in syllabification is determining to which syllable the consonants of a consonant cluster belong. Here the principle mentioned above helps determining the parsing of clusters. For instance, according to Kenstowicz (1994: 251), a word-internal cluster like /tr/ is parsed as an onset in English on the basis of its sonority profile, of the MOP, according to which as many consonants as possible are placed in the onset if they comply to the SSP, and also on the basis of the appearance of this cluster word-initially (word edges = syllable edges principle). A cluster like /tl/, on the other hand, is parsed as a coda-onset sequence notwithstanding its rising sonority profile because /tl/ is not a sequence found word-initially. Alternatively, these latter clusters have been analysed as bogus clusters in Government Phonology (Harris 1994: 182ff), clusters that qualify neither for branching onsets nor for coda-onset sequences and are separated by an empty nucleus. The bogus cluster option is an analytical tool that is not exploited further here, since the existence of this third kind of cluster is grounded in the very syllabification principles that are being questioned in this contribution: the sonority sequencing principle, on which governing relations are based (although restated by means of segmental complexity), and the word edges = syllable edges principle,

according to which word-internal clusters may be granted tautosyllabicity only if they also appear word-initially. Clusters that are qualified as bogus in the literature display a phonological behaviour consistent with a coda-onset cluster but cannot be considered as such on theory-internal grounds: the cluster does not appear word-initially (i.e. /tl/) or it cannot establish a coda-onset governing relation because of its segmental content (i.e. /tr/), i.e. its sonority profile. Accordingly, only the tautosyllabic and the heterosyllabic parse are considered as options in the remainder of the discussion.

2.1   Predictions of current syllable theory

The expected syllabification of clusters on the basis of the principles mentioned is as follows: typically an obstruent/liquid sequence (*muta cum liquida)* makes a well-formed branching onset on the grounds of its optimal sonority profile. It is therefore tautosyllabic according to the MOP, unless the language disallows sequences of consonants word-initially or complex onsets. Obstruent/nasal sequences can be parsed differently depending on the presence of such clusters word-initially. Other consonant clusters form heterosyllabic sequences by default, on the ground of their sonority profiles, flat in the case of obstruent/obstruent and sonorant/sonorant, falling in the case of sonorant/obstruent. The heterosyllabic parse of these clusters word-internally is generally substantiated by phonological evidence cross linguistically. When these clusters are found word-initially, where a tautosyllabic cluster is predicted, on the other hand, a number of adjustments are made to the theory in order to separate the consonants in the cluster so that the syllable at hand can be described as complying to the SSP, also at the beginning of the word. A case in point are sC clusters (cf. Goad 2011 for a summary).

2.2   Diagnostics for syllabification

A troublesome by-product of the view shaped by the current perspective on the syllable and sonority, where clusters with sonority rising clusters are predicted to be tautosyllabic and word edges are predicted to coincide with syllable edges is that, accordingly, obstruent-sonorant clusters are often described in the literature as tautosyllabic by default, without any phonological evidence brought to bear (Bhatt & Nikiema 2006: 95 among others). Likewise, word-initial clusters are, in many contributions, labelled onsets, regardless of the sonority sequence profile and without examining the phonological behaviour of the cluster (Blevins 1995: 218–219; Morelli 1999; Proctor 2009: 121; Kreitmann 2012, among others). The resulting situation is that different types of clusters are often granted tautosyllabic status

if word-initial, and that conversely tautosyllabic status is granted to word-internal obstruent/sonorant clusters.[1]

In this contribution I draw attention to the fact that empirical data, introduced and discussed in Section 3, show the possibility for all kinds of clusters to display both tautosyllabic and heterosyllabic behaviour and that, in addition, word-initial and word-internal clusters may not always display the same syllabic parse in the same language. Accordingly, a pre-established parse of clusters cannot be maintained. On the contrary, syllabification ought to be established for each cluster and each position and each language on the basis of phonological evidence, i.e. processes like open syllable lengthening/closed syllable shortening, vowel differentiation in open and closed syllables, syllable-driven allophonies or through syllable weight revealed by poetic meter. Another helpful diagnostic of syllabification is the process of vocalic epenthesis in clusters, since a cluster triggering vowel insertion hints at its phonotactic ill-formedness or 'looseness'[2] and helps revealing the syllabic parse of the consonants.

3.   Syllabification of clusters

This section presents and discusses empirical data concerning the syllabification of clusters with different sonority profiles in different positions of the word.

3.1   Obstruent/sonorant syllabification

Data from several languages show that the same cluster may display a different syllabification word-initially and word-internally. More specifically, in this section obstruent/sonorant clusters (*muta cum liquida)* are considered. In a number of languages, these clusters typically display heterosyllabic behaviour word-internally but tautosyllabic behaviour word-initially. This is the case for the reconstructed Indo-European language (Byrd 2015), Icelandic (Gussmann 2002: 173), Ilokano (Hayes & Abad 1989), Munster Irish (Green 2003), Lithuanian (Steriade 1997), Sanskrit (Cooper 2015) among others. To exemplify in detail all these languages

---

1. It is sometimes the case that the labels 'onset' and 'cluster' are used with different meanings in the literature. To avoid any ambiguity, in this contribution the label 'onset' refers to a prevocalic consonant or to a prevocalic consonant cluster that phonologically behaves as a single consonant. A cluster on the other hand is any sequence of consonants, regardless of its syllabic parse, which is established on phonological evidence.

2. Looseness in this case means that, for instance, anaptyxis in an obstruent/sonorant cluster helps indicating that the cluster is not bound by a branching onset relationship.

would by far exceed the scope of this paper. The argument is merely exemplified with data from the Italo-Romance dialect of San Valentino in Abruzzo Citeriore (henceforth San Valentino see Passino & Pescarini 2018). In this dialect of the Upper-Southern dialect group, spoken in Eastern Abruzzo, the evolution from Latin shows identical evolution of tonic vowels before *muta cum liquida* and in closed syllables, as opposed to open syllables. This is shown in (1) by means of the evolution of A boldfaced (cf. Savoia 2015: 116 for other examples of heterosyllabic *muta cum liquida* in Italo-Romance)[3]:

(1) Evolution of A in the dialect of San Valentino:
     open syllable   closed syllable   before *muta cum liquida*
     CASA(M) > ˈkəsə STAB(U)LA(M) > ˈstallə MACRU(M) > ˈ**m**agrə

The syllable-driven allophony concerning /a/ is also observed in synchrony, as shown by the adaptation of Italian loanwords, elicited from an informant and reported in (2):

(2) open syllable   closed syllable before *muta cum liquida*
     [fer.ˈrə.rə] < Ferrari [ˈ**f**aks] < fax   [ˈ**m**atrə] < Matra
      (car brand)   (car name of the Simca brand)

I consider this sufficient evidence to indicate a heterosyllabic parse of this cluster in word-internal position. This notwithstanding, *muta cum liquida* clusters abound in word-initial position. Here, their tautosyllabic status is guaranteed by the fact that, as opposed to sonorant/obstruent and obstruent clusters respectively arising from apheresis or present in loanwords, they do not undergo any repair in the language, as shown in (3):

(3) Word-initial clusters in Sanvalentinese
     Tautosyllabic (unrepaired)   Heterosyllabic (repaired)
     ˈtrəvə <TRAVE(M) 'beam'   rət'tokə < (U)RTICA(M) 'nettle'
                                  p**i**siˈkoləkə, cf. It. *psychologue*
                                  'psychologue'

The data presented so far show that identical clusters in word-initial and word-internal positions may display different syllabic parses, as opposed to what standard syllable theory would predict. Another prediction of standard syllable theory that is not borne out, if one considers the same data, is that *muta cum liquida*, being of rising sonority, are parsed tautosyllabically.

    The possibility of a heterosyllabic parse of *muta cum liquida* clusters, however, is not restricted to the word-internal position, but can also be recorded in

---

3. Some exceptions exist due to the fact that the dialect is situated in a transitional area. The heterosyllabic parse of *muta cum liquida* is more consistent in the area south of San Valentino.

word-initial position. A language that illustrates this case is Moroccan Arabic, where phonological and phonetic evidence, as well as evidence coming from versification, hints at the absence of complex onsets in the language (Cantineau 1946; Kaye et al. 1986; Dell & Elmedlaoui 2002; among others).

In conclusion, obstruent/sonorant clusters may display a heterosyllabic parse both in word-medial and word-initial position. Arguably, the latter case only concerns languages not allowing complex onsets. The former, on the other hand, is documented in languages like Sanvalentinese, where a tautosyllabic parse must be assumed for the word-initial position.

3.2   Obstruent cluster syllabification

A sequence of obstruents displays a flat sonority profile and the prediction of standard syllable theory is therefore that it is parsed heterosyllabically. Languages that comply with this prediction can be easily found in Romance, among others. Here Standard Italian data are used to exemplify the argument, since the language offers diagnostics for heterosyllabic syllabification of obstruent clusters both in word-initial and in word-internal position. Let us begin by examining the allomorphy of the masculine definite article, a diagnostic of cluster syllabification in word-initial position: the allomorph *il* is selected before tautosyllabic clusters whereas *lo* is selected before heterosyllabic clusters or instrinsic geminates (Marotta 1993). In (4) it is shown that the article selected before obstruent clusters is *lo*, thus guaranteeing for their heterosyllabic parse:

(4)   Allomorphy of the masculine definite article in Italian

il: before simple Cs and tautosyllabic clusters	*lo*: before heterosyllabic clusters and intrinsic geminates
il topo 'the mouse'	lo pneumatico 'the tyre
il treno 'the train'	lo pterodattilo 'the pterodactyle'
il pjatto 'the dish'	lo stato 'the state'
il clone 'the clone'	lo ʃʃame 'the swarm'

The classic diagnostic of syllabic parse in word-internal position for Italian, on the other hand, is provided by tonic lengthening, according to which stressed vowels in open non-final syllables are long, while those in closed syllables are short. Unfortunately, in the literature on the topic (see Fava & Magno Caldognetto 1976), I was not able to find measurements of tonic vowels before obstruent clusters, since those clusters are non native and only appear in few loanwords. Preliminary measurements carried out by Ulfsbjorninn (2017), however, show that vowels are short before those clusters, thus hinting at their heterosyllabic status. To strengthen the argument, we can appeal to another diagnostic used by Ulfsbjorninn in

the same work, although with a slightly different aim, namely the application of *Gorgia Toscana* (Giannelli & Savoia 1978; Marotta 2008), a spirantisation process typical of Tuscan Italian. *Gorgia Toscana* applies to singleton intervocalic stops and to stops involved in complex onsets, as shown in (5a). The non application of *Gorgia Toscana* to the first member of obstruent clusters exemplified in (5b) hints at the heterosyllabic parse of such clusters[4]:

(5) a. /apre/ [ˈaɸre]
     b. /ipsilon/ [ˈipsilon] *[ˈiɸsilon]

Notwithstanding the situation just exemplified, empirical data showing a tautosyllabic parse of obstruent clusters are also available. I discuss here the Gallo-Italic dialect of Bologna, where a number of word-initial clusters different from *muta cum liquida* obtained from a syncope process. Of these 'sonority offending' clusters only sonorant/obstruent clusters were repaired by epenthesis. The obstruent clusters were accepted in word-initial position without repair and pattern with tautosyllabic *muta cum liquida*, as shown in (6). This situation was first pointed out by Loporcaro (1998) on data of Grizzana Morandi, a village near Bologna.

(6) Word-initial prosthesis in Bolognese

	No prosthesis		prosthesis
muta cum liquida		obstruent clusters	sonorant/obstruent clusters
CRUCEM > ˈkrauz		VESTIRE > ˈftir	NEPOTEM > aŋ ˈvaud
'cross'		'to dress'	'nephew'
TRES > ˈtri:		VESSICA> ˈpsiga	LAETAMEN > aldam
'three'		'bladder'	'manure'

Synchronically, in addition, the patterning of obstruent clusters with tautosyllabic *muta cum liquida* can be observed by looking to vowel/zero alternation in *sandhi*. As shown in (7), no epenthetic vowel surfaces between the monoconsonantal preposition *d* 'of' and words beginning with obstruent clusters and *muta cum liquida*. Only the nucleus preceding the preposition vocalises its previous nucleus and both clusters behave like single consonants, hinting at the tautosyllabic parse of the clusters.[5]

---

4. The author reaches the conclusion that the cluster is a bogus cluster. Having excluded this analytical tool from the possible analyses in 2, I take vowel length as a diagnostic of the presence of a syllable closed by the first member of the cluster.

5. In this contribution the definition 'tautosyllabic cluster' refers to all clusters behaving like single consonants. Accordingly this definition also include cases in which the consonants occupy only one structural slot (monopositional clusters).

(7)   Bolognese
      ˈpke   'sins'   na moʧa d pke   'a lot of sins'
      ˈkraida   'clay'   na moʧa d kraida   'a lot of clay'

As for the word-internal situation, distributional evidence can be used as a diagnostic. Obstruent clusters, as opposed to heterosyllabic clusters and similarly to *muta cum liquida*, may appear after sonorant codas, as shown in (8):

(8)   ˈmandla 'almond'   ˈmandga 'sleeve'

The data examined showed that obstruent clusters in these dialect also display tautosyllabic behaviour in the word-internal position. The Bolognese data examined so far showed that obstruent clusters may display a uniform tautosyllabic parse, as opposed to what standard theory would predict. Another prediction of standard syllable theory that is not borne out is uniform syllabification across word positions. Although in Bolognese the tautosyllabic status is confirmed across positions, in the Emilian dialect spoken in Reggio Emilia a difference exists between the syllabification in utterance-initial position and word-initial position in *sandhi*, which is not expected by standard assumptions. The situation, whereby obstruent clusters are free to stand in utterance initial position but trigger prosthesis when preceded by a consonant, is illustrated in (9):

(9)   Microvariation in Emilian epenthesis patterns

Utterance initial position		Utterance-internal, word initial position	
No initial prosthesis		*Initial prosthesis*	*No initial prosthesis*
Bolognese	Reggiano	Reggiano	Bolognese
pke	pke	na moʧa d æpke	na moʧa d pke
'sins'	'sins'	'a lot of sins'	'a lot of sins'

As well as *muta cum liquida*, obstruent clusters show syllabic behaviour that is not predicted by standard syllabic assumptions.

3.3   Sonorant/obstruent syllabification

A sonorant/obstruent cluster has a reverse sonority profile and it is thus predicted to be parsed heterosyllabically. This is indeed the situation generally found cross-linguistically. If we use the diagnostic provided by open syllable lengthening, we find that languages where this process is active do not display lengthening before sonorant/obstruent sequences. This is shown in (10) with data from the Apulian dialects of Ruvo di Puglia and Andria (Savoia 2015: 75, 90), where respectively long vowels and diphthongs surface in open syllables and short vowels appear in closed syllables:

(10) *Ruvo di Puglia* [ˈnuːvə] 'new'   [ˈkurtə] 'short'
     Andria   [ˈkraitə] '(I) believe' [ˈleŋgwə] 'tongue'

In this contribution however, we discuss evidence of an alternative behaviour of such clusters word-internally, as well as evidence setting apart liquid/obstruent and nasal/obstruent clusters. In the diachrony of Emilian and Piedmontese dialects, vowels preceding liquid/obstruent clusters, as well as the ones preceding *muta cum liquida,* yielded open syllable outcomes, as opposed to nasal/obstruent clusters and geminates, as shown in (11) exemplifying the evolution of A in Bolognese (Coco 1970):

(11) Outcomes of A in Bolognese (Emilian)
     Open Syllable = ε   Closed syllable = a
     LA.CU(M) > lεːg 'lake'   GAL.LU(M) > gaːl 'cock'
     CA.PRA(M) > kεːvra 'goat'   CAT.TU(M) > gaːt 'cat'
     A.LTU(M) > εːlt 'tall'   SAN.GUE(M) saŋgv 'blood'
     BA.RCA(M) >bεːrk 'boat'

Other dialects like the Lombard variety of Villa di Chiavenna even show a subtle difference in the parsing of NC clusters: Nasals followed by voiceless obstruents leave the preceding syllable open while nasals followed by voiced obstruents do not, as in (12):

(12) Outcomes of A in Villa di Chiavenna (Savoia 2015: 177)
     Open Syllable = ε   Closed syllable = a
     ˈsε.nε 'sain.F'   ˈkan.nε 'canne'
     ˈfε.ntʃ 'enfant.M'   ˈgam.bε 'jambe'

This means that, notwithstanding the overwhelming presence of data showing sonorant/obstruent heterosyllabicity, also the tautosyllabic parse must be allowed for, as a genuine possibility of human language. 'Tautosyllabic parse' of a cluster is here intended as a cluster not displaying a heterosyllabic behaviour. It is not implied that a sonorant/obstruent cluster is a complex onset. The tautosyllabic surface effect may depend on the ability of the voiceless obstruent to govern the rhyme (Savoia 2015: 178) or on the ability of sonorants, in some perspectives allowing for the presence of empty nuclei separating the consonants in the clusters (Lowenstamm 1996; Scheer 2004), to branch on the following nucleus (Pöchtrager 2001). By the same standards, a word-initial sonorant/obstruent cluster could in principle display a heterosyllabic parse, that in standard analysis corresponds to extrametricality of the first consonant of the cluster, or a 'tautosyllabic' parse, whereby the sonorant is syllabic or where it shows a phonological behaviour different from the one displayed by heterosyllabic clusters. I compare here nasal/obstruent clusters in word-initial and word-internal position using again fieldwork-data

on Sanvalentinese. These data, illustrated in (13), show that, while nasal-obstruent clusters in Sanvalentinese are clearly heterosyllabic in word-internal position (the preceding syllable displays closed syllable allophones), in word-initial position the situation is different. Word-initial nasal/obstruent clusters, as well as other clusters, result from the process of apheresis targeting word-initial unstressed vowels in the language. While liquid/obstruent clusters are repaired by epenthesis, nasal/obstruent clusters and geminates are not. Accordingly, these latter pattern with *muta cum liquida* clusters word-initially displaying a tautosyllabic behaviour, while the former pattern with word-initial obstruent clusters obtained from loanwords being heterosyllabic and repaired:

(13)   Different parse of nasal/obstruent clusters in Abruzzese (San Valentino dialect).
Word-internal.
Heterosyllabic (closed syllable allophone).
kw**a**ndə < QUANTUM cf. open syllable allophone of /a/: kəsə < CASAM.

Word-initial

Tautosyllabic (unrepaired)	Heterosyllabic (repaired)
trəvə <TRAVE(M) 'beam'	rəttokə <(U)RTICA(M) 'nettle'
mbutə < cf. It. *imbuto* 'funnel'	pisikoləkə cf. It. *psicologo*
tʃitə <OCCIDERE 'to kill'	psychologue'

As was the case for other kinds of clusters, also sonorant/obstruent clusters display a phonological behaviour that is not expected by standard syllabic assumptions.

3.4   Sonorant cluster syllabification

Clusters of sonorants are rarer than other clusters in all positions, as they are frequently repaired. They are described as heterosyllabic and the repairs have been analysed as driven by sonority requirements holding also between coda-onset sequences: due to the absence of sonority steep, bad syllable contact is established in the cluster (Venneman 1988; Davis 1998, among many others). The brief examination of the phonology of these clusters allowed by the space constraints on this contribution is enough to ground one of the main reflections underlying this contribution, namely that sonority, as it is commonly understood, is not enough to account for all the attested pattern of syllabification recorded cross-linguistically. Only by considering the word-initial position, clusters of sonorants display a different syllabification, as shown in (14) again with data of Bolognese prosthesis of initial clusters arisen from syncope:

(14) Word-initial prosthesis in Bolognese
      No prosthesis   Prosthesis
      mlaˈranza 'orange'   alˈɲam 'wood'

This evidence, among other, suggests that the major class features are not enough to predict syllabification (cf. also the different NC parse shown in (13)) and that if sonority is helpful, a finely grained scale must be allowed for.

4.   Final remarks

In this contribution some data have been discussed in order to point out the difficulties that the principles upon which the standard theory of syllabification is based find in coping with the extent of the documented empirical data documenting:

1. Different syllabic parses of the same cluster in a given language
2. Unpredicted syllabic parses are documented cross-linguistically

While this contribution has perhaps offered some systematization of the data, a number of the problems discussed were known. The reaction to this is generally to put effort in showing that the data that cannot be captured by standard sonority-based assumptions are peripheral and that the principles governing syllabification are at work even in languages that seem to violate them (Spencer 1996: 91; Riad 2004: 187, among others) or to diminish the strength of the MOP as a heuristic (Hayes 2009 among others) and of SSP from Principle to Preference (Blevins 1995: 211; Loporcaro 1998, among others). Strong versions of the principles are maintained through added theoretical machinery such as coda-capture, extrametricality, extraprosodicity, and different kinds of adjunctions. The data discussed, however, if taken seriously, may open interesting paths of research and lead ultimately to a theory of the syllable able to account for all attested patterns, regardless of their frequency.

References

Bhatt, Parth & Nikiema, Emmanuel. 2006. In *The Structure of Creole Words*, Parth Bhatt & Ingo Plag (eds), *85–106*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110891683
Byrd, Anthony. 2015. *Reconciling Indo-European Syllabification*. Leiden: Brill.
Blevins, Juliette. 1995. The syllable in phonological theory. In *Handbook of Phonological Theory*, John Goldsmith (ed.) , 206–244. Oxford: Basil Blackwell.
Cantineau, Jean. 1946. *Les Parlers arabes du Hôrân: Notions générales, grammaires*. Paris: C. Klincksieck.

Coco, Francesco. 1970. *Il dialetto di Bologna*. Bologna: Forni.

Cooper, Adam I. 2015. *Reconciling Indo-European Syllabification*. Leiden: Brill.

Davis, Stuart. 1998. Syllable contact in Optimality Theory. *Korean Journal of Linguistics* 23: 181–211.

Dell, François & Elmedlaoui, Mohamed. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-010-0279-0

Fava, Elisabetta & Magno Caldognetto, Emanuela. 1976. Studio sperimentale delle caratteristiche elettroacustiche delle vocali toniche e atone in bisillabi italiani. In *Studi di Fonetica e Fonologia*, Raffaele Simone Ugo Vignuzzi & Giulianella Ruggiero (eds) , 35–79. Roma: Bulzoni.

Giannelli, Luciano & Savoia, Leonardo M. 1978. L'indebolimento consonantico in Toscana. *Rivista Italiana di Dialettologia* 2: 25–58.

Goad, Heather. 2011. The representation of sC clusters. In *The Blackwell Companion to Phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds), 898–923, Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444335262.wbctp0038

Green, Antony Dubach. 2003. Extrasyllabic consonants and onset wellformedness. In *The Syllable in Optimality Theory*, Caroline Féry, Ruben van de Vijver (eds), 238–253. Cambridge: CUP. https://doi.org/10.1017/CBO9780511497926.010

Harris, John. 1994. *English sound structure*. Oxford: Blackwell

Hayes, Bruce. 2009. *Introductory Phonology*. Oxford. Blackwell.

Hayes, Bruce & Abad, May. 1989. Reduplication and syllabification in Ilokano. *Lingua* 77: 331–374. https://doi.org/10.1016/0024-3841(89)90044-2

Jespersen, Otto. 1904. *Lehrbuch der Phonetik*. Leipzig: B.G. Teubner.

Kaye, Jonathan, Echchadli, Malika & El Ayachi, Souad. 1986. Les formes verbales de l'arabe marocain. *Revue Québecoise de Linguistique* 16(1): 61–98. https://doi.org/10.7202/602580ar

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kreitmann, Rina. 2012. On the relations between [sonorant] and [voice]. In *Consonant Clusters and Structural Complexity*, Philip Hoole, Lasse Bombien, Marianne Pouplier Christine Mooshammer Barbara Kühnert, (eds), 33–70. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614510772.33

Loporcaro, Michele. 1998. Syllable structure and sonority sequencing: Evidence from Emilian. In *Romance Linguistics: Theoretical Perspectives, Selected papers from the 27th Linguistic Symposium on Romance Languages [Current Issues in Linguistic Theory 160]*, Armin Schwegler, Bernard Tranel & Miriam Uribe-Etxebarria (eds), 155–170. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.160.13lop

Lowenstamm, Jean 1996. CV as the only syllable type. In *Current Trends in Phonology. Models and Methods*, Jacques Durand & Bernard Laks (eds), 419–441. Salford, Manchester: ESRI.

Marotta, Giovanna. 1993. *Selezione dell'articolo e sillaba in italiano: un'interazione totale? Studi di grammatica italiana, XV*,255–296.

Marotta, Giovanna. 2008. Lenition in Tuscan Italian (Gorgia Toscana). In *Lenition and Fortition*, Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds), 235–271. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211443.2.235

Morelli, Frida. 1999. The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory. PhD dissertation, University of Maryland.

Parker, Stephen. 2002. Quantifying the Sonority Hierarchy. PhD dissertation, University of Massachusetts Amherst.

Passino, Diana & Diego Pescarini. 2008. Il sistema vocalico del dialetto alto-meridionale di San Valentino in Abruzzo Citeriore con particolare riferimento agli esiti di Ū. Atti del XXVIII Congresso internazionale di linguistica e filologia romanza (Roma, 18-23 luglio 2016). edited by, Antonelli, Roberto, Videsott, Paul & Glessgen, Martin. BiLiRo 15 Société de Linguistique Romane / Strasbourg: Éditions de linguistique et de philologie. 484–497.

Proctor, Michael Ian. 2009. Gestural Characterization of a Phonological Class: The Liquids. PhD dissertation, University of Yale.

Pöchtraeger, Markus. 2001. Finnish Consonant Gradation. MA thesis, University of Vienna.

Riad, Tomas. 2004. Syllabification and word division in Gothic. *Journal of Germanic Linguistics* 16(2): 173–202. https://doi.org/10.1017/S1470542704000388

Savoia, Leonardo M. 2015. *I dialetti italiani*. Pisa: Pacini.

Scheer, Tobias. 2004. *A Lateral Theory of Phonology*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110908336

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In *Language Sound Structure*, Mark Aronoff & Richard T. Oehrle (eds), 107–136. Cambridge MA: The MIT Press.

Sievers, Eduard. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf & Härtel.

Spencer, Andrew. 1996. *Phonology*. Oxford: Blackwell.

Steriade, Donca. 1982. Greek Prosodies and the Nature of Syllabification. PhD dissertation, MIT.

Steriade, Donca. 1997. *Phonetics in Phonology: The Case of Laryngeal Neutralization*. Ms, UCLA.

Ulfsbjorninn, Shanti. 2017. Bogus clusters and lenition in Tuscan Italian: Implications for the theory of sonority. In *Sonic Signatures*, Geoff Lindsey & Andrew Nevins (eds), 278–296. Amsterdam: John Benjamins. https://doi.org/10.1075/lfab.14.c16

Venneman, Theo. 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Guyter.

Language in context

Diachronic and synchronic lexical interactions in the Italo-Balkan linguistic space

From Latin *lucanĭca* to Italo-Albanian *lëkëngë*

Francesco Altimari
University of Calabria

This contribution presents a case study in the intricate diachronic and synchronic interrelations between languages in the Italian and Balkan linguistic spaces. Words related to Latin *lucanica* ('type of sausage') are investigated and hypotheses on their pattern of diffusion are surveyed and discussed. Among other points, the contribution touches on the etymology of *lucanica* itself, the intermediation role of Byzantine Greek, the multiple outcomes of the word in the Romanian group, its present-day distribution in Italo-Romance and Arbëresh dialects.

Keywords: Albanian dialectology, Balkan linguistic space, food lexicon, lexical diffusion, loanwords

1.   Introduction

A recent work of mine considers a word in the Albanian food lexicon (*tumacë* 'home-made pasta') which before landing in Italy with the Albanian people had a long and adventurous linguistic voyage, from the Caucasus to the Apennines (Altimari 2018). I noted that trying to reconstruct the linguistic contact between populations from a word is a difficult scientific endeavour, even if not an impossible one. This is true in some lexical domains, such as cooking, which are among the most hybridized, considering the long timespan in which the history of foods develops within the history of the people who produce and consume them. The hybridization level increases when we deal with one of the linguistic domains, which are contaminated *par excellence* by exchanges and relations among peoples – and one of these is represented precisely by the Balkan peninsula. In the challenge posed by the conditions just described, we emblematically chose to widen our investigation by considering a word belonging to the food lexicon once again.

The choice fell on *lucanica* (Lat.), *λουκάνικον* (Gr.), *lugàniga* (North It.) 'sausage type' – which we intend to analyze both diachronically and synchronically, taking into account its distribution both in the Balkan and in the Italian linguistic space as well as the interactions which we have been able to track between these two linguistic domains, on the basis of historically attested occurrences of the lexeme recorded in the various languages considered. We agree with Guigoni (2006), who in an interesting essay on the food culture of the Mediterranean, stresses that human food habits represent a long term phenomenon, tendentially conservative as to the products used, but also with respect to the cooking techniques employed and the consumer styles, destined to be perpetuated for many generations down the ages. Nor should we forget the central role played by Braudel's triad formed by grain, vine and olive, which is associated with all Mediterranean civilizations up to our days, together with cheese and pork meat. Precisely pork meat will be at the center of this study following the history of *lucanica* through the traces of it present in the various languages being examined.

2.   Attestations of the lemma in the Italic and Balkan linguistic space

Let us begin with the Italic linguistic space. In the *Treccani* on-line dictionary, at the lemma *lugàniga* (or *lugànega*), Lat. *lucanǐca* (lit. 'Lucanian sausage'), we read: "North Italian, widespread in Veneto and Lombardy, synonym with sausage in general, and local name for some traditional types of sausages of those regions, most of them shaped as long and thin cylinders, without the characteristic links. The Italianized forms *lucànica* and *lugànica* are also used".[1] In the *Dizionario Etimologico della Lingua Italiana* (Cortellazzo & Zolli 2011), the same lemma *luganiga*, meaning 'typical sausage of Veneto and Lombardy', is associated with the following first attestations, for the different regional varieties recorded in different North Italian Romance varieties: *luganega* (Lombardy) 1315, *luganica* (Veneto) 1477, *lucanica* (Lombardy) 1550, *luganiga* (Veneto) 1621.

The *Tesoro della Lingua Italiana delle Origini* (Leonardi 1997), confirms the first Italo-Romance attestation of *luganega* 'typical sausage of North Italy' in the Lombardy area in 1315. Documents in the Capitoline Historical Archive prove that *locaniche* were consumed in Rome by the Jewish community (we surmise in their kosher version), in the XVI-XVII century (Stow 1988 quoted in Weingarten 2011). The Macedo-Rumanian dictionary of Papahagi (1974) quotes the Italianized form *locànica* from the *Vocabolarietto anconitano-italiano* (Spotti 1929); according to the author, this designates a sausage used only by Jews.

---

1. English translations from the originals are provided by the editors.

Another classical etymological Italian dictionary (Battisti and Alessio 1975) confirms that *lugàniga* 'type of sausage' is attested in the Northern Italo-Romance area: Old Venetian *lugànega* (1300), Old Friulian *lugània* (1288), modern Friulian *lujanie*. In any event, it is a term "well represented in modern Northern and Alpine dialects". The origin is taken to be Latin *lūcānica* 'sausage of Lucania', first attested in *De lingua latina* by Marcus Terentius Varro (116–27 BCE), where in Book V, cap. XXII, 7 we read: *quod fartum intestinum e crassundiis, Lucanicam dicunt, quod milites didicerint* ("a sausage made with the large intestine of pork is called *lucanica* because the soldiers learned how to make it from the Lucanians"). The *Lexicon philologicum* (Martini 1655)[2] reports other Latin sources of the Classical and Late period in which *lucanica* is mentioned: Cicero,[3] Apicius,[4] Martial[5] and Isidore of Seville.[6]

Let us now consider the distribution of the lemma in the Balkan linguistic space. According to Battisti and Alessio (1975), supported also by Karanastasi, the passage from Latin *lūcānicum* to Medieval Greek *lukānikón* happened around the IV century, with consequent attestations in the Balkan linguistic space.[7]

---

2. Martini illustrates the distribution of the lemma in French and German: "Gloss. Lucanica ἀλλαντία vel Lucanica ἀλλάντια ut sit diminitivum. Gallis est andouille vel endouille, puto ab ἔνδον intus, ut intestinum ἐνδόθιον, iisdem est saulsisse, an saulse condimentum, quasi salsum, an ἀλλᾶς ab ἅλς, quod dulce Chald". (Martini 1655: 217).

3. Cicero (106–43 BCE) mentions *lucanica* in the *Epistulae* (*Ad familiares*), specifically in the *Epistula* 9: "Solebam antea delectari oleis et lucanicis tuis".

4. Apicius (circa I century BCE – I century CE) deals with the *lucanica* in the Book II, IV of *De re coquinaria*: "Lucanicas similiter ut supra scriptum est: Lucanicarum confectio teritur piper, cuminum, satureia, ruta, petroselinum, condimentum, bacae lauri, liquamen, et admiscetur pulpa bene tunsa ita ut denuo bene cum ipso subtrito fricetur. Cum liquamine admixto, pipere integro et abundanti pinguedine et nucleis inicies in intestinum perquam tenuatim perductum, et sic ad fumum suspenditur".

5. Martial (39/40–104 CE) quotes *lucanica* in two books of his *Epigrammaton Liber*: lib. IV, 46 (*et Lucanica uentre cum Falisco*) and lib. XIII, 35 (*Filia Picenae venio Lucanica porcae: pultibus hinc niveis grata corona datur*). In both quotations, the *lucanica* is associated with food traditions and cultures of ancient central Italian peoples, such as Falisci and Piceni.

6. Isidore of Seville (circa 560–636 CE), in his work *Etymologiarum sive Originum libri XX*, quotes *lucanica*, but does not stand out for his originality, simply reproducing previous etimologies: "Lucanicae dictae, quia prius in Lucania factae sunt. Farcimen, caro concisa et minuta, quod eo intestinum farciatur hoc est, impleatur cum aliarum rerum commixtione", (lib. XX, 2, 28).

7. According to most scholars, the first attestation of Greek *loukanika* appears in a papirus that describes the travel to Antioch of an Egyptian dignitary called Teofane, who keeps a detailed record of the expenses sustained, including also six purchases of *loukanika* (Matthews 2006).

Nevertheless, other sources seem to bring forward this first presence in Greek texts by a couple of centuries, given an attestation in a papyrus datable between the II and III century CE.[8] For the Roman and Old Byzantine phases the dictionary by Sophocles (1887), which covers the period from 146 CE until 1100, attests λουκάνικον (as a synonym of ἀλλᾶς 'sausage'). Similarly, λουκάνικο(ν) is present in the dictionary of Kriaràs (1988), in sixteen volumes, comprising popular medieval Greek attested from 1100 to 1669.

Through the mediation of Byzantine Greek the lemma would then have been transmitted to the languages already then in contact with Medieval Greek. Until the VI-VII century these were represented solely by Proto-Albanian variaties and by Eastern Romance varieties (Romanian and Macedo-Romanian); subsequently, the various Slavic varieties spoken by communities living in the Balkans after the VI century were added. As a consequence of the transmission of Medieval Greek λουκάνικον, Battisti and Alessio's (1975) dictionary attests *rucanicò* 'sausage' in the Grecanic of Calabria and specifically in Bovese, as well as Albanian *lëkonkë, -ongë*, Macedo-Romanian *lucangăi*, Bulgarian *lukanka*.[9]

This overall picture, presented by Battisti and Alessio's dictionary, should be refined with respect to Aromanian and Vlach varieties. Papahagi (1974), which is to this day the most comprehensive dictionary of Aromanian (or Macedo-Romanian), records the lemma *lucáncu ~ lucánic* with the meaning of *cîrnat* 'sausage', generically (bur erroneously!) indicating *lëkonkë* as its Albanian counterpart (Papahagi 1974: 749). In fact, it is not possible to make reference either to common Albanian or even to Albanian dialectal varieties of the Balkan area, includig Arvanitic ones[10] present in the insular and peninsular territories of the Greek Republic. Rather this lemma has left no traces in Balkan varieties of Albanian, while it is abundantly (though not systematically) present in Italo-Albanian.

---

8. My heartfelt thanks go to the friend and colleague Katerina Papatheou for her detailed indications on the presence of λουκάνικον in Greek texts, from the Ellenistic to the Neo-greek era. The papirus source quoted in the text is *SB*.26.16553, 8: χανιχων(χανικῶν, or [ῷῶν] χανικῶν, or [λου]κανικῶν) (δραχμαὶ), and can be referred to a period comprised between 175–225 CE. According to K. Papatheou it is the most ancient one. One would therefore anticipate by almost two centuries the first Greek attestation, hitherto considered canonical and referred to the IV-V century CE: Ὀζόστομος λουκάνικον ὀπτῶν καὶ ἐπὶ πολὺ προσφυσῶν, κυνέαν αὐτὸ ἀπειργάσατο ('a man with bad breath was cooking sausages, but he breathed on them so much that he made them become manure'), *Philogelos* (*Φιλόγελως*) (Dawe 2000).

9. The form *lokanka* is attested as a borrowing into Slavic languages in Miklosich (1886).

10. We have not found any traces of *lëkëngë* in any of the Arvanite language areas, which have been subjected to investigation by the renowned Albanian and Balkan scholar Titos Jochalas (cf. his important linguistic monographs on the Arvanita in Greece: Jochalas 2000, 2002, 2006, 2011).

Capidan (1922), in his in depth study devoted to the linguistic relations between Albanian and Aromanian, also record several dialectal Aromanian variants of this term, always with the meaning of *cârnaţ* 'sausage': *lucanicu, lucáncu, lucangu*, bu also *culucancu*.[11] Similarly, in his etymological dictionary, de Cihac (1879) reproduces the variant *culucancu* under the lemma *lucánică* 'espèce de saucisse' specifying that it is a term not popular in Daco-Romanian, but widespread in Macedo-Romanian.

In a more strictly defined Romanian domain, in the *Dicţionarul Limbii Române* (DLR, 2008) we find *lucánică* with the meaning of *carnât* 'sausage', *caltaboş* and *sângerete* 'blood sausage'.[12] Here however the lemma is not considered a Latin borrowing from late antiquity, mediated by Greek, or a more recent borrowing from some Romance variety of Italian area, but is rather recorded as a Latinism, attested *in dicţionarele din trecut* in dictionaries of the past'. Indeed it is found in a few dictionaries of the 19th century quoted by the DLR. In the Buda Lexicon (Royal Hungarian University 1825) we find "*Lucanica* f.pl.e. *Лѹкáникα* V. Cárnaţu (sausage)". The Latinizing academic dictionary of Laurian and Massim (1876) reports "*Lucanica*, s.f., lucanica, botulus cruore distentus, carnaciu grossu implutu cu *sange*, de unde se dice si sangerete, *sangeretecu*, pre candu *lucanica* in limb'a latina classica, semnifica mai multu carnaciu de friptu, care se face d'in carne; lucanica ferta cu mória; lucanic'a este buccata grea pentrù stomachu; (de la *Lucan*, poporu italicu)".

As already mentioned, the DLR traces the origin of the term, present in the whole of the Romanian linguistic area, back to Latin *lucanica*. By contrast, in his monograph on Romanian suffixes, having pointed out that in Romanian only few words of Greek origin end with the suffix -ic, Pascu identifies lucanic e lucangu as borrowings from the Greek λουκάνικον (Pascu 1916: 396).[13]

---

11. In Capidan (1922: 526), *culucancu* is explained by following Meyer (1895: 551) as a result of contamination with the Albanian word, *kolé* "sausage made out of liver, chopped meat and herbs". In the academic dictionary of present-day Albanian (Instituti i Gjuhësisë dhe i Letërsisë 2002: 591) *kolé,~ja* corresponds to *Zorrë bagëtish të imëta, e mbushur me mëlçi, me mish të grirë e me erëza, që hahet zakonisht e pjekur, suxhuk* "gut of a sheep or a goat stuffed with liver, minced meat and herbs, which is usually eaten roasted, synonymous with *suxhuk*".

12. The word, however, is reported here with a different accentuation: *lucaníca* in some sources (S.P. Barcianu, W. Alexi) and *lucănică* in others (D. Pontbriant, R.M. Gheţie).

13. "in -*ic* se termină şi câteva substantive, ale căror radicale sunt ngr. [...] *Lucánic, lucangu* ,cârnaţ' (Dal) < *λουκάνικον*". It is worth noticing that these (reciprocally synonymous) Romanian lexical items, also attested in Aromenian, have close analogoues in Albanian, which presents an older *lëkëngë* in the Arbëresh in Italy next to a more recent one *llukanìk* in Balkan Albanian.

We leave aside for the time being the question of the origin or source of this borrowing in the two old Balkan languages Romanian and Albanian; for, the question is rather complex and difficult to decide, if one cannot make reference to attestations in published or written texts, in any event associated with an established date. The source or origin could have been Greek for both language (hypothesis A). But one could also hypothesize a passage from Greek to Albanian, with a mediation of the latter, via Aromanian varieties, to attain Romanian (hypothesis B).[14] Nor can we exclude the possibility of a passage from Greek to Aromanian, after which (via the mediation of the Macedo-Romanian variety) the term would have passed into Romanian and Albanian (hypothesis C).

What we can exclude is that these borrowings are to be considered synonims, to be imputed to the same phase in the history of these languages, given that Romanian *lucangu*, Aromanian *lucáncu* and Arbëresh *lëkëngë* are synchronically interrelated and certainly older than Romanian *lucanic*, Aromanian *lucánic* and Albanian *llukanìk*, which are equally synchronically interrelated. Consider hypothesis B. The stress shift from the second to the third syllable would lead us to conlude in favour of an expected adaptation to the stress rules of the host language 1, namely Albanian, passed on to the host language 2, namely Romanian. Suppose on the other hand we adopt hypothesis C. The host language 2, in this instance Albanian, would have adopted the Aromanian lexical item *lucánic*, with stress shift on the penultimate syllable, preferred in this language, where it is nowadays attested as *llukanikë* and no longer as *lëkëngë*,[15] as was the case in the Medieval period. The latter has been maintained in the Arbërisht area, following the rules of

---

14. We find this hypothesis in Meyer-Lübke (1911: 370), where as to the lexeme n. 5134 *lucanica* 'Art Wurst', the Macedo-Romanian *lucanic*, *lucangu* is assumed to derive from the Albanian *lukaník, lẹkongẹ*.

15. The only Albanian dictionary in the Balkan Area to attest it is the one by Thimi Mitko (1820–1890), prepared after the second half of the 19th century and remained unpublished to our days. The merit goes to Qirjazi (2014) for retrieving and printing this important lexicographic repertoire and producing a valuable philological edition (Mitko 2014). On page 312 we find *lëkëënk-a* 'sausage', with [flj.6, 238] as a source. At first sight I had assumed the abbreviation to refer to De Rada's *Fjamuri Arbërit*, but a direct check on number 6 of the same newspaper did not confirm my hypothesis. The doubt was solved by Dhori Qiriazi himself (p.c.), who hypothesizes that flj. might actually not be the abbreviation of *Flamur* (which recurs elsewhere in Mitko), but a reference to *Fylétia e Arbenoré* by Demetrio Camarda (1867), even though the reference in Mitko (flj.6, 238) happens not to match. Considering the close ties that Mitko entertained with the most noteworthy exponents of the Arbëresh cultural scene, G. De Rada and D. Camarda among them, in this case, *lëkënka* can certainly be interpreted as a lexical arbëreshism borrowed from the Italo – Albanian lexicon by some of his interlocutors' writings. In the Greek-Albanian part of his dictionary, Mitko also reports (p. 559) "λουκάνικον = αιματίτης. kurma-gják-u (πρμτ.) kollofáce-ja καὶ kolē –ja (Χᾱν). Τὸ δε kollofáce είναι εκ χοιρίνου κρέατος. [A939]."

phonetic adaptation of Greek *loukanikón* into Albanian and preserving the position of the main stress.

There is another possible conclusion that we may put forth here: the likely derivation both of Albanian *llukanikë* and of Romanian (and Aromanian) *lucánic* or *lucanic* not from λουκάνικον but from another Greek lemma, namely λουκανίτικος 'fat' (referred to sausages). For both Albanian and Romanian, we could then indicate a date after the XVI century for the possible entry of these lemmas in their lexicons, by direct contact with Byzantine Greek. Indeed λουκανίτικος is first attested in a Byzantine satire in the first decades of the XVI century (Eidener 1977).[16] The position of stress at least in Albanian would coincide with that of the Greek source and we can easily follow the passages in the phonetic adaptation from the source language to the target language. <ι> in post-tonic position reduces to schwa /ə/ of; /τ/ in the word-final consonantal sequence /τ/ + /κ/ is subsequently deleted, given that in the meantime, in the morphological assimilation of the Greek loanword into Albanian, the final –ος suffix was lost.

3. History and current distribution of the lemma

Let us now try to follow the events that have affected our lemma in its linguistic history between the Italian peninsula and the Balkan peninsula from its first attestation, which as we already mentioned, in found in Varro's *De lingua Latina* (I century BCE), where *lucanica* 'Lucanian sausage' is attested for the first time.

According to an hypothesis which seems to us objectively more plausible, from the Latin stream would have originated the Greek stream, first attested around the III century CE; the first attestation would be Late Classical Greek λουκάνικον, λουκανικόν, derived from Lat. *Lucanicum* and attested in a papyrus datable to the II-III century CE (see fn 8). In late antiquity, as we already saw, other attestations of λουκανικόν are subsequently found in the Mediterranean area under Greek influence. In the Middle Ages (and later) the same term is attested by the most authoritative lexicographical repertories concerned with Late Classical Greek and with Byzantine Greek, already briefly reviewed – to then arrive to those of the Modern age, up to today. Among the latter we mention those of Andriotis (1995 [1951]: 191) and Triantaphyllidis (1998).

Needless to say, given the continuity in the presence of λουκανικόν in all Greek speaking areas, from the Hellenistic age to today, attestations of it could not

---

16. The German Byzantine scholar Eideneier publishes this Byzantine satire, which had already appeared in Venice in 1553. He edited two other versions based on the known manuscripts of this work and dating back to respectively 1515–1519 and 1540.

be lacking in those South Italian areas, such as Salento and Calabria,[17] characterized in the Early Middle Ages by a strong Byzantine presence and by a notable Greek cultural heritage.

From the double matrix, the Latin one with *lucanicum* and the Greek one with *λουκανικόν*, two different linguistic traditions would then originate in the spreading of the term to a Mediterranean and European context. The Italic one, of Latin origin, is attested today in North Italian dialects and is probably at the basis of loanwords in various Western linguistic varieties (French, Spanish, Portuguese, Basque). The Balkan one, of Medieval Greek origin, has spread by contact to the neighbouring linguistic areas (Albanian, Aromanian, Romanian and Bugarian), adapting phonetically to those languages.

The mass migrations that have affected Greek communities in the Early Middle Age and Albanian communities in the Late Middle Age have meant that *lukānikón* has come back from the Balkan area into the South Italian areas where it possibly originates. Thus it is found in Bovese in the form *rucanicò*[18] – chronologically more ancient – and in Arbëresh in the form *lëkëngë* – Chronologically more recent. These are the regions that border with historical Lucania, which at the times of the Latin domination, presented wider borders than the Lucania or Basilicata of today, comprising also parts of Southern Campania (the Vallo di Diano) and parts of Northern Calabria.

Paradoxically, there is no linguistic trace of *lucanica* in South Italian Romance varieties today – but, according to the linguistic documents that have survived from late antiquity and the Middle Ages, there hasn't been even in past centuries. This could be explained by the non-necessity for a community to auto-denominate a product not special, but common with its ethnonym, i.e. with a brand-name which was recognized *extra moenia*, outside of Lucania. This term in South

---

17. The 'Greekness' of *lukānikó* could be another critical point, though not a diriment element, in the old linguistic quarrel on how to interpret Bovesia's Greekness. In the Rohlfsian theory, it is to be regarded as a permanence of a Magno-Greek archaicism. In the Alessi-Battisti-Parlangeli theory, it is rather a medieval trait resulting from the Byzantine domination. Lastly, Karanastasis' theory integrates an assumed autochthonous Magno-Greek Greekness (*à la* Rohlfs), with subsequent additions from the Hellenistic koiné, reinvigorated by the linguistic pressure of the Byzantine domination. Therefore, this lemma could also be used either in a conservative sense (*rukanikò* being the only attestation of the Latin type *lucanicum* in Southern Italy, although mediated by the Greek *lukanicon*) or in an innovative sense (the 'return' among Lucanians of a rather archaic Grecism).

18. On the meaning of *λουκάνικο* in Bovesian Greek, cf. Karanastasis (1991: 361–2): "Ἄν δό κρέα τοῦ κουννίου κάν-νουσι τὰ ρουκανικὰ κιόλα τὲ φ-φουρίνε καὶ τὴν ἀρτυσία = ἀπὸ τὸ κρέας τοῦ γουρουνιοῦ κάνουν τὰ λουκάνικα, ἐπίσης τὶς φορίνες καὶ τὴν ἀρτυσιὰ (τὸ λίπος). Τὰ ρουκανικὰ ἔναι κρεμαμένα".

Italy nowdays lives on linguistically thanks to the Greeks and Albanian, through the loanword *loukanikon*, of Greek origin but of Latin descent, brought back during the Byzantine dominion over South Italy (V-XI century). It is also continued through the loanword *lëkëngë*, taken along by Albanians at the time of their great migrations into the regions then belonging to the Kingdom of Naples, starting from the XV century.

As second paradox is that we find this loanword in the best part of Italo-Albanian varieties, attested as *lëkëngë~likëngë* to designate 'sausage', namely in Calabria,[19] in Molise,[20] in Apulia[21] and in Sicily[22] – except that *lucanica* is lacking precisely in the Albanian varieties of the Lucanian area. We find no trace of it, since instead of *lëkëngë*, the Italo-Romance loanword *saucicë* is attested both in the Albanian-speaking are of the Vulture river (Barile, Maschito e Ginestra),[23] and in the Albanian-speaking area of the Pollino mountains, represented by the dialects of San Paolo Albanese and San Costantino Albanese.[24] The latter, located at the heart of the so-called Lausberg area (Martino 1991), present special and adiosyncratic linguistic characters, including the preservation of many archaic forms, both in the grammar and the lexicon. This makes the study of these varieties of old Tosk Albanian very interesting both for dialectology (Albanian and Romance) and for the history of the Albanian language, given the historical linguistic contact between these two archaic languages.

---

19. *Lëkëngë* is attested in the dialect of Frascineto (Giordano 1963: 227). It is widespread, however, in the Albanian-speaking area of Calabrian Pollino: Shundi (2016), in the recent *E(t)nogastronomia arbëreshe*, reports *likëngë ~ lëkëngë* 'lloj salsiçeje, type of sausage', for the communities of Civita, Firmo, Acquaformosa, San Basile and Lungro, as well as in Piana degli Albanesi. Also in Calabrian Arbëria, the word is attested in Spezzano Albanese, but neither in the Albanian varieties of the Sila, nor in the Valle Crati, the Crotonese or Catanzarese regions.

20. *likëngë* 'sausage' is attested in Molisan Albanian (Pignoli & Tartaglione 2007: 98; Fiorilli 2002: 71).

21. *lëkëngë* 'sausage' is attested in the dialect of Casalvecchio di Puglia (Massaro 2010: 84).

22. We find that *lëkënkë~likënkë* 'sausage' is already included as a Sicilian-Arbëresh term in the small dictionary by Chetta (1763: 214). It is attested in the dialect of Piana (Gerbino 2007: 60): *lëkëngë ~ likëngë* 'sausage'.

23. This piece of information has been offered to me by my friend Donato Mazzeo, native of Barile and familiar with the Albanian spoken in the Vulture area. I thank him heartily for his precious help.

24. *salsiccia*: *saucìc-saucica- saucica-saucicat* "pork meat finely chopped, seasoned with salt, grounded red pepper and wild fennel, encased in the smaller guts of the very same animal" (Scutari 1991: 235).

Beside the Albanian varieties spoken in Lucania, there is no trace of *lucanica* even in Lucanian Romance dialects. In the excellent lexicographical survey of the dialectal varieties of Basilicata edited by Bigalke (1980: 732), the term does not appear in any of the Lucanian dialects spoken today. Instead we find [saw'tsiʧʧ] with its variants, displaying the extreme fragmentation of the dialects of the Southern Lucanian area, with some phonetic outcomes taking us back to a rather archaic linguistic phase: [saw'tsəʧʧ], [saw'tsitsts], [sau'tsiʧʧ], [savi'tsitsts], [saßə'tstsiʧ], [saß'siʧʧ], [saßə'sic][25] and so on.

In principle another hypothesis could be supported by the truly strange absence from Southern Italo-Romance varieties of any foodstuff associated with the Latin term *lucanica*. The linguistic diachronic continuity from Latin to Romance (which remains to be explained) is indeed strangely circumscribed to a restricted dialectal area of the country, in practice the Northern are, especially the Lombard-Venetan area, far removed from the geographical origin area, Lucania, which would have given birth both to the food item and to its name. The survival of the sole Lombardy-Veneto *luganega* could find some explanation, if we hypothized the preservation of the food stuff in the later Balkan area, together with the loanword *lucanicum* of Latin origin, mediated by the Greek *lukānikón*.

Since late antiquity *lucanica* had a considerable gastronomic fortune in the Mediterranean-Balkan context. Here, as often happens in the domain of food, further elaborations are determined by new ethnic, social, economic and alimentary host contexts. These elaborations turn an imported food into a food perceived as typical and proprietary in the common perception – and no longer as extraneous and taken from other traditions (let's not forget the intense contacts in commerce and food in existence between Venice and the East and vice versa in the XII-XVIII century). The gastronomic fortune mentioned above could have led the foodstuff to be experted together with its name, through the Medieval Greek *loukanikón*, back into its orginal area of attestation, namely the Italian peninsula, though only in the Northern regions and not in the Southern ones.

This would be supported by the outcomes *luganega* oe *luganiga* (with the corresponding Italianized forms *lucanega* and *lucanica*), due to the usual application of the phonetic rules of voicing of intervocalic *k* to *g* characteristic Northern Italian variety of Lombardy and Veneto.[26] Those terms would have designated the typical sausage imported by the Venetians from the Mediterranean areas of the

---

25. Phonetic transcriptions have been made by translating the alphabet adopted by the Romance scholars listed in Bigalke's Lucanian dictionary to IPA, for ease of use and because this system is now by far the most widespread and well-known.

26. Cf. the sections 197 (intervocalic *k* in Northern Italy) and 198 (intervocalic *k* in Southern Italy) in the first volume of Rohlfs's precious historical grammar (Rohlfs 1966: 269–270).

Balkans and would have been circumscribed to the Northern Italian dialectal areas (where it established itself from the XIV century on) for geopolical reasons, namely because of the domination of those regions by Venice.

If we accept this interpretation, we could even not give too much credit to Varro's etymology. It could be a popular etymology that could have led the author astray, and would have then justified the spreading of the etymology itself, conditioning the explanation that we give of the term up to this day. Evidently, the benefits and delicacies of food culture in the past (even without having big distribution chains like McDonald's at one's disposal) were propagated among the different peoples even then, without paying excise. Globalization was already in place, in the Mediterranean domain as well, already in those far away times. People's emigrations, with their cultures (including food) and their languages, in centuries closer to us have then brought us back closer together, creating the complex linguistic situation well exemplified by the diachronic comings and goings of our word in the Italo-Balkan linguistic space.

We have tried to get out of the quagmire by proposing possible and plausible interpretations, reaching a truth which at the end of a long and difficult (though never wasted) research, is nevertheless provisional and falsifiable, as we are reminded by Popper. Our solutions, hopefully creditable and convincing, endeavor to understand the trajectories followed by the peoples and their languages.

At the end of this linguistic investigation around an ancient word like *lucanica/luganega/λουκάνικον/rukanikò/lucáncu/lëkëngë*, in our research voyage on the food lexicon of the Italo-Balkan linguistic space, we have in any event marked some significant stages. We are pleased that these not casually come to coincide with some of the many regional varieties and minority languages investigated by Leonardo Savoia, in his fruitful and original research between Italy and the Balkans. We celebrate him today at the end of a productive, prestigious and brilliant academic career, thanking him deeply for what he has generously given and taught us.

Ad multos annos! Για πολλά χρόνια! De mai mulți ani! Për shumë vjet!

Acknowledgment

References

Altimari, Francesco. 2018. Italian-Balkan linguistic interactions in the Italian-Albanian food lexicon: A short history of Tumacë 'Home-made Pasta' from the Caucasus to the Apennines. In *Balkan and South Slavic Enclaves in Italy: Languages, Dialects and Identities*, Thede Kahl, Iliana Krapova & Giuseppina Turano (eds), 2–31. Newcastle upon Tyne: Cambridge Scholars.

Andriotis, Nikòlaos P. 1951[1995]. *Ετυμολογικό λεξικό της κοινής νεοελληνικής*, 3rd rev. edn. Thessaloniki: Idryma Manoli Triantafyllidi.

Battisti, Carlo & Alessio, Giovanni. 1975. *Dizionario Etimologico Italiano*, 5 Vols. Florence: G. Barbera.

Bigalke, Rainer. 1980. *Dizionario dialettale della Basilicata*. Heidelberg: Carl-Winter-Universitätsverlag.

Capidan, Theodor. 1922. Raporturile albano-române. In *Dacoromania – Buletinul 'Muzeului Limbei Române'* 2: 444–554.

Dawe, Roger D. (ed.) 2000. *Philogelos*. München: K.G. Saur.

De Cihac, Alexandre. 1879. *Dictionnaire d'Étymologie Daco-Romane. Éléments slaves, magyars, turcs, grecs-moderne et albanais*. Frankfurt: Ludolphe St-Goar.

DLR = Dicționarul Limbii Române. 2008. New series, Vol. 5, Letter "L". Bucharest: Editura Academiei Române.

Eideneier, Hans. 1977. *Spanos: Eine byzantinische Satire in der Form einer Parodie*. Berlin: Walter De Gruyter.

Chetta, Nicolò. 2008. *Leksiko liti kthiellë arbërisht.* (1763) Palermo: Besa.

Cortelazzo, Manlio & Zolli, Paolo. 2011. Il nuovo Etimologico, *DELI-Dizionario Etimologico della Lingua Italiana*. Bologna: Zanichelli.

Fiorilli, Giuseppe. 2002. *Dizionario arbëresh di lemmi, frasi e motti in uso nella tradizione di Ururi*. Lanciano: publisher unknown.

Gerbino, Gaetano. 2007. *Fjalori arbërisht-italisht i Horës së Arbëreshëvet – Dizionario arbëresh-italiano della parlata di Piana degli Albanesi*. Palermo: publisher unknown.

Giordano, Emanuele. 1963. *Fjalor i Arbëreshvet t'Italisë – Dizionario degli Albanesi d'Italia*. Bari: Edizioni Paoline.

Guigoni, Alessandra. 2006. L'alimentazione mediterranea tra locale e globale, passato e presente. In *Saperi e sapori del mediterraneo*, Alessandra Guigoni & Radhouan Ben Amara (eds), 81–92. Cagliari: AM&D Edizioni.

Instituti i Gjuhësisë dhe i Letërsisë. 2002. *Fjalor i shqipes së sotme, botim i dytë i ripunuar (me rreth 35.000 fjalë)*. Tirana: Botimet Toena.

Jochalas, Titos P. 2000. *Άνδρος. Αρβανίτες και Αρβανίτικα*. Athens: Pataki.

Jochalas, Titos P. 2002. *Ευβοια. Τα Αρβανίτικα*. Athens: Pataki.

Jochalas, Titos P. 2006. *Ύδρα. Λησμονημένη γλώσσα* (2 vols). Athens: Pataki.

Jochalas, Titos P. 2011. *Η Αρβανίτικα στο Μοριά, χρονικάπορείας, τόμ Α΄-Β΄*. Athens: Pataki.

Karanastasis, Anastasios. 1984–1992. *Ιστορικόν λεξικόν των Ελληνικών ιδιωμάτων της Κάτω Ιταλίας. Αθήνα: Ακαδημία Αθηνών*, Vols. I–5: vol. I (1984), vol. II (1986), vol. III (1988), vol. IV (1991), vol. V (1992).

Kriarás, Emmanouìl. 1988. *Λεξικό της Μεσαιωνικής Ελληνικής Δημώδους Γραμματείας 1100–1669 του Εμμανουήλ Κριαρά, Τόμος Ι΄* (vol IX). Thessaloniki: Typ. Sfakianakis.

Laurian, August T. & Massim, Ion C. 1876. *Dictionariulu Limbei Romane*, Vol. 2. Bucharest: Noua Typographia a Laboratoriloru Romani.

Leonardi, Lino (ed.). 1997. Tesoro della Lingua Italiana delle Origini (TLIO). Florence: Istituto del Consiglio Nazionale delle Ricerche. <http://tlio.ovi.cnr.it/TLIO/>

Martini, Matthia. 1655. *Lexicon philologicum, praecipue etymologicum et sacrum, in quo latinae et a latinis auctoribus usurpatae tum purae tum barbarae voces ex originibus declarantur, etc.* Frankfurt: Thomas M. Goetzelnil.

Martino, Paolo. 1991. *L'area Lausberg isolamento e arcaicità*. Rome: Sapienza University of Rome.

Massaro, Mario. 2010. *Dizionario comparato delle parlate arbëreshe di Casalvecchio e Chieuti*. Apricena: Malatesta.

Mathiesen, Johan. 1993. The children of Lucanica. *Petits Propos Culinaires* 43: 62–63.

Matthews, John. 2006. *The Journey of Theophanes. Travel. Business, and Daily Life in the Roman East*. New Haven CT: Yale University Press.
https://doi.org/10.12987/yale/9780300108989.001.0001

Meyer, Gustav. 1895. *Albanesische Studien, IV: Das griechisch-südrumänisch-albanesische Wortverzeichniss des Kavalliotis herausgegeben und erklärt* . Wien: F. Tempsky.

Meyer-Lübke, Wilhelm. 1911. *Romanische Etymologisches Wörterbuch*. Heidelberg: Carl-Winter-Universitätsbuchhandlung.

Mitko, Thimi. 2014. *Fjalëtore shqip-gërqisht/gërqisht-shqip & ndihmesa të tjera leksikografike*. Tirana: Academy of Sciences of Albania.

Miklosich, Franz. 1886. *Etymologisches Wörterbuch der Slavischen Sprachen*. Wien: Wilhelm Braumüller.

Papahagi, Tache. 1974. *Dicționarul dialectului aromân, general și etimologic, ediția a doua augmentată*. Dictionnaire Aroumain (Macédo-Roumanin), général et étymologique, 2nd edn, enlarged. Bucharest: Editura Academiei Republicii Socialiste România.

Pascu, Giorge. 1916. *Sufixele Românești*. Bucharest: Edițiunea Academiei Române.

Pignoli, Maria Luisa & Tartaglione, Guido. 2007. *Dizionario albanese molisano (Parlate di Portocannone e Ururi)*. Cosenza: University of Calabria.

Rohlfs, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti*, Fonetica. Turin: Einaudi.

Royal Hungarian University. 1825. *Lexicon Valachico-Latino-Hungarico-Germanicum a pluribus auctoribus decursu triginta et amplius annorum elaboratum est*. 1825. Buda: Royal Hungarian University.

Scutari, Pasquale. 1991. *Spigolature lessicali sulla parlata degli Arbëreshë di S. Costantino Albanese*. Scanzano Jonico: publisher unknown.

Sophocles, Evangelinus Apostolides. 1887. *Greek Lexicon of the Roman and Byzantine periods (From B.C. 146 to A.D. 1100)*. New York NY: Charles Scribner's Sons.

Shundi, Andrea. 2016. *E(t)nogastronomia arbëreshe: Shtjellohen 610 receta gjellësh dhe ëmbëlsirash*. Tirana: Botime Vllamasi.

Spotti, Luigi. 1929. *Vocabolarietto anconitano-italiano*. Genève: Leo S.Olschki.

Stow, Kenneth. 1988. Life and society in the community of Rome in the sixteenth century. *Pe'amim* 37: 55–66.

Triantaphyllidis, Manolis A. 1998. Λεξικό της κοινής νεοελληνικής. Thessaloniki: Ìdryma Manoli Triantafyllidi.

Weingarten, Susan. 2011. Ancient Jewish sausages. In *Cured, Fermented and Smoked Foods. Proceedings of the Oxford Symposium on Food and Cookery*, Helen Saberi (ed.), 365–375. Devon: Prospect Books.

CHAPTER 23

Lexical-semantic analysis of the political language

Studies between 1960 and 1980

Paola Desideri

D'Annunzio University of Chieti–Pescara

The aim of this paper is to provide an overview of the studies on political language produced between 1960 and 1980. This is a very important period for research in this area, as it marks a significant development of lexical-semantic analysis concerning linguistic corpora, which are essential to the history of Europe. The speeches by the protagonists of the French Revolution, the totalitarian languages, the messages by last century parties and leaders, as well as the dictionaries and different political lexicon adopted by East and West Germany are notably the favoured subjects for semantic in-depth analysis, to understand the significant connotations and meaning effects.

Keywords: lexicon, semantics, political language

1.   Aspects and trends

The most flourishing development of the studies on political language lies in the lexical-semantic area, which turns out to be the largest one. Mostly between 1960 and 1980 there was an important exploration of political *corpora,* considered to be extremely representative and paradigmatic to the history of Europe. Thanks to a boost in research in fields like lexicology and analysis of utterances seen as verbal *performances*, namely speeches, mainly conducted in the French area,[1] allocutions uttered by protagonists of the French Revolution were mostly taken into consideration, as well as messages produced by dictators, leaders and the 20th century parties, and political linguistic codes, deliberately different as to their meaning, adopted by the Deutsche Demokratische Republik (DDR) and Bundesrepublik Deutschland (BRD).

---

1. Fundamental scientific reference for lexicology and lexical analysis has been Dubois's work (1969).

© 2018 John Benjamins Publishing Company

It is a rather complex operation to develop even a non-exhaustive overview of the most interesting works on semantics related to the political language,[2] both because of their significant amount and above all because of their inherent interdisciplinarity with other similar survey fields. In fact, morphosyntactic and rhetorical argumentative considerations are often rightly coexistent in such studies.

The lexical-semantic analyses of the period concerned, in the methodological variety which recognises and distinguishes them, set the following fundamental aims:

1. evaluate the meanings of specific political terms in ideologically significant discursive formations and then identify their semantic, contextual value;
2. explain the connotative power of watchwords, keywords, clichés, euphemisms, neologisms, foreignisms, circulating under socially problematic circumstances and periods;
3. indicate, according to the needs of structuralism, the function, structure and role performed by some discursive elements in the textual construction;
4. reconstruct the diachrony of politically significant expressions and formulas from the first uses onwards;
5. provide quantitative data, percentages, and frequency indexes for specific language uses;
6. finally classify and make an inventory of the socio-political locutions adopted in specific situations, developing lexical repertoires and dictionaries.[3]

Consequently, for a proper description of studies most explored[4] by the scholars who dealt with political language from 1960 to the late 1970s, we believe it is suitable to group such analyses according to the thematic macro-categories considered, highlighting the most used criteria as well as operational methods. As the practice of political communication finds its raison d'être into the ideological and social formations, into the balance of power, within historical circumstances, it is obvious that the interest of lexicologists, lexicographers and semanticists were particularly interested in texts related to events which were traumatic and decisive for society. Those texts regard speeches made by charismatic chiefs and

---

2. On this rich branch, a bibliographic index as well as a comment on the most representative contributions are found in Desideri & Marcarino (1980: 59–78, 26–33). As for the origins and development of research on political language, see Desideri (2009).

3. As concerns the true political dictionaries, see Cofrancesco (1978), who identifies three different sorts of dictionaries: the ideologists, the utopians, and the problematics.

4. This occurs for two main reasons: on one hand, the long tradition this linguistic discipline has, on the other hand, the exploration since the 1980s of new, prolific fields such as the pragmatic linguistics, the enunciative theory, the speech acts, the typology of speeches.

leaders, popularized by political parties and movements and set in particularly turbulent times.

The manipulative activity of political communication is expressed in different ways. In the word level, a leading role is played by watchwords, whose vague content can be 'filled' with distinct meanings, sometimes even antithetical depending on the points of view. If frequently repeated, they have a strong unifying power, also being the basis for the creation of slogans and stereotypes. In this respect, Dieckmann (1969: 101–106)[5] is indeed of the opinion that the watchword is a phenomenon of *parole* and not of *langue* and, consequently, it does not absolutely exist, but receives its appellative emotional function only in connection with specific contextual conditions, so differentiating itself from buzzwords.

An important role is also played by keywords, namely the most used denominations to indicate groups, states, policies, ideals and problems (Zimmermann 1969: 63). They are marked by a strong social penetration and produce a positive or negative evaluative effect on the events in the discursive context (e.g. *nation*, *homeland*, *population*, etc.). Manipulation is also achieved by the cliché, which is a tendentious simplification easy to memorise, a manageable, reassuring term or formula that reduces much more complex and intricate realities to a common denominator. Such procedures, typical of verbal, political stereotypy, are deployed more than we think and studied in detail by Roegele (1972), Quasthoff (1973) and Harnisch (1974).

2.   The French Revolution language and beyond

The first important international meeting on political lexicology, devoted to *Formation et aspects du vocabulaire politique français, XVIIᵉ -XXᵉsiècles*,[6] was ar-

---

5. Dieckmann's considerations concerning the political language's peculiarities represent one of the most influential reference points on the subject. The West German research of those years took inspiration from them. Dieckmann, proceeding on the basis of Lasswell's categories relating to the language of political propaganda (*credenda, miranda, formula*), seeks to specify the nature of the ideological, institutional language, and finally the technical one of the administrative areas (Dieckmann 1969: 47–52). On the essential lesson by Lasswell, founding father of *Content Analysis* applied to the political language and mass communications, see Desideri (2009: 47–48). Specifically for ideologically polysemous words Dieckmann stops to examine the meanings and uses of an extremely multi-purpose term: *Demokratic* (Dieckmann 1969: 70–75). On the equivalent Italian lemma, *democrazia*, see the semantic and diachronic, rigorous observations by De Mauro (1958).

6. The Conference Proceedings were published in three monographic issues of *Cahiers de lexicologie* 13 (1968), 14 (1969), 15 (1969).

ranged by the *Centre de lexicologie et textes politiques* – a laboratory created in 1967 as a research team of the École Normale Supérieure of Saint-Cloud – and it was held in the same place from 26 to 28 April 1968.[7] The copious lines of lexicological research on linguistic political *corpora* become systematic in France, and debute with the discourse analysis and first lexicometrical attempts. Saint-Cloud is, in fact, the place where political lexicology is made official as an independent scientific discipline. Still in the sixties the équipe *Révolution Française* is also constituted, composed of eminent researchers, such as Annie Geffroy, Jacques Guilhaumou, Régine Robin, the latter manager of the "Centre d'Analyse du discours de l'Université de Lille III".[8]

Most studies on the French Revolution terminology and texts, and above all on the idiolect of some remarkable protagonists of that period, undoubtedly one of the most important in modern age, have been carried out as part of this équipe's activity. There are two main directions of such research:

a. the systematic elaboration, made by means of a mechanised perforation and statistical calculations, of alphabetical indexes, frequency lists, keywords placed in well-defined contexts and allocutions as those of Hébert, Robespierre, Roux and Leclerc (Geffroy 1973, 1974, 1980);
b. the examination of how given semantic areas as "anarchie", "féodalité", "noblesse" worked and the identification of ideological routes of meaning in a series of texts, such as the famous newspaper-pamphlet *Père Duchesne* by Hébert (Guilhaumou 1975, 1980) and *Courrier des départements* by Gorsas (Renaudie 1975, Ruetsch 1975: 117–134), without excluding the songs made between 1789 and 1794 (Mancuso 1975) that were examined according to the enunciative procedures in terms of both contents and typology.

In the same directions the following collective works also move: Robin (1973); Guilhaumou, Maldidier, Prost & Robin (1974); Drigeard, Geffroy, Laplace, Tournier & Wagner (1977). However, such works are just one part of the prolific study on the language and semantic organisation of revolutionary and pre-revolutionary texts. For obvious reasons of space we have to exclude the abundant *mémoires de maîtrise* and the doctoral theses presented from 1969 on, which tended to focus on the semantic procedures of the revolutionary forces' discourse (Jacobins, Montagnards, Sansculottes), the distinctive features of Robespierre and Saint-Just's lexicon, the utopian socialists' lexicon, the expressive peculiarities of

---

7. It should be noted that the Colloque of Saint-Cloud took place just a few days before students and workers began the French May.

8. In this regard see the rich monographic issue of the *Bulletin du Centre d'Analyse du discours de l'Université de Lille III* 2 (1975) on the subject *Sur la révolution française*.

the well-known *Cahiers de doléances,* which were some registers in which the people's complaints were noted.

With the French Revolution, namely when it became necessary to create precise currents of opinion, the actual political propaganda starts and, consequently, political language is also born as a subcode for the debate and comparison of positions. If in the early 18th century politics is still on the static, conceptual positions of a science separated from the contingent real-life pressures, in the late 18th century the variable and opened power management imposes a new face to the same politics, which becomes a technique of the action and conditioning of citizens aiming at an immediate result.

Along with French scholars, also Italian and German ones have dwelt on this tumultuous time, developing specific semantic analysis from different perspectives: we particularly mention Krauss (1970b) who examines the development of the meaning of *révolution* and Leso (1976) who analyses the diachronic process and semantic implications of the term *moderato,* especially after the Jacobin experience. Let us not forget the Swedish scholar von Proschwitz (1966), who looks in the political lexicon before and after the revolution.

The entry of revolutionary gallicisms into the other languages circuit and their adoption in the persuasive processes of political marketing are investigated by: Dieckmann (1964), who observes the presence of stylistic features typical of the French Revolution in the German political propaganda; Zolli (1964,1965), who aims at analysing the gallicisms entered in the Italian political language in the 18th century; Leso (1970) was interested in the meanings and contexts of *cittadino* foreignism during the three years of revolution; Medici (1976) was attracted to the Italian Jacobin period and Carofiglio (1975), who looks into Marat's political language with its connotative features and meaning effects.

We also have to mention Barny's essay (1978) and Blanchard's volume (1980), particularly interesting for the sharpness and originality of survey, who examines, according to a three-pronged approach (stylistic, semantic and semiological), the nature of the revolutionary expression and linguistic registers used, the spaces of elocutive representation (podium, scene) and the functioning of texts (pamphlets, orations).

In Blanchard's analysis (1980: 69–105) prevails Saint-Just's, Vergniaud's and Robespierre's speeches; of which Blanchard identified the most significant lexical forms, denominations and appellatives addressed to powerful people and the most effective tropes, while the meaning paths operating under very complex communication practices were underlined, such as the famous theatrical performances and revolution parties.

Keeping with revolutionary *corpus* analysis, from the methodological point of view, a quite isolated case is the aforementioned essay by Slakta (1971) on the

Cahiers de doléances, analysed following the criteria of Fillmore's generative se-
mantics. Not even the linguistic appearance of the pre- and post-revolution de-
cades has been overlooked. In fact, as regards the political language of the "Ancien
Régime", Krauss (1970a) and Siccardo (1978) examine the semantic values and
connotations of *patriotism*, of which an excellent lexicographical repertoire
is proposed.

As for the 19th century language, the indispensable volume by Dubois (1962)
accounts for the innumerable lexical-semantic phenomena that are typical of the
French sociopolitical vocabulary until 1872, seen in its different evolutionary stag-
es: of the latter, useful diachronic data are given, contents and oppositions are
rebuilt and lastly, interferences and signification ambiguities are highlighted.

Finally, Vardar (1973) takes into account a specific historical cross section,
the Restoration period, and mainly focusing on a number of speeches by lead-
ing figures representing different political views, he explains the semantic fields
where the political subcode from 1815 to 1830 is operative, provided with a rich
lexical filing.

3.   The totalitarian languages and the discursive styles of leaders and party movements

Certainly, the twelve years passed between the first Saint-Cloud's meeting in 1968
and the second one about *Lexicologie politique du français moderne*[9] – always in
the same seat from 15 to 20 September 1980 – have brought the birth and growth
in France of a wide range of studies about the terminological particularities as
well as various lexical-semantic aspects that characterized the broadest and most
different political discursive formations. We have already mentioned from the
revolutionary lexicon to the confederal one, from the communist to the socialist
production, from De Gaulle, Giscard d'Estaing and Mitterand's idiolects to the
significant connotations and mass-media transformations.

The lexical and stylistic characteristics of the speeches of the most historical
figures have always attracted the attention of academics, who were committed to
measuring the ideological importance of some lemmas, proving neologisms and

---

9. The congressional section which some scholars have joined mainly, clear interest to some
arguments, was dedicated to *Vocabulaire du texte politique et syndical*, with relative subsections:
Discourse Communiste (1927–1946); *Presse contemporaine (1960–1980)*; *Révolution française*;
Vocabulaire syndical (1919–1979); *Discours nazionalistes contemporains.* Some papers disclosed
to this Colloque will be published some years later on *Mots* (2, 3, 4, 5 1981–1982) and *Cahiers
de lexicologie* (40, 41 1982).

euphemisms meaning effects as well as considering the semantic power of persuasive locutions and verbal manipulations. There has always been an incisive terminological selection suitable to the circumstances and events that have "seduced" public opinion; as a consequence, it has influenced judgments and behaviors.

The *parole* of the dictators Mussolini, Hitler and Franco has been subject of remarkable analytic interest with its respective historical and linguistic codifications (Fascism, National Socialism, Francoism language), in order to understand the hidden mechanisms of signification, the magic of lexicon, the power of the totalitarian languages' slogans and code words, through which entire generations gained experience.[10]

For the semantic and stylistic elements of both the Mussolinian and Fascist regime language, we count a valuable contribution from Leso (1973); Lazzari (1975); Leso, Cortelazzo, Paccagnella & Foresti (1977) who focused their attention on the most significant rethorical procedures inside the texts; finally Simonini (1978).

About the constitutive features of the Nazi and Hitler production[11] we must remember von Polenz (1967) on the policy of the linguistic purism used by national socialism; Lange (1968); Bork (1970); Winckler (1970) about *Mein Kampf*; Enzi (1971) who realized a real dictionary about the markedly aggressive lemmas adopted by Nazism. As for the historical periods before and after Nazi Germany, we mention Cobet (1973) about anti-Semitism vocabulary during Bismarck's period; Mueller (1973: 24–42) on the Third Reich and ex-DDR's neologisms; Bergsdorf (1978) about the national socialist and communist language.

Regarding Franco's language, not particularly studied during the period between 1960 and 1980, we send back to Cillán Apalategui (1970) and Rebollo Torío (1975, 1978).

The linguists' attention is raised not only by the particular lexical profile of the major European dictators, but also by that of other important political figures of the ancient and modern context: they focused on the semantic procedures, the stylistic features and the semic nucleus at the basis of the discursive production of the charismatic leaders. Regarding persuasive forms and stylistic procedures of Lenin's speeches, we send back to Glimm's essay (1961), but we must mention the fundamental research conducted by Šklovskij, Tyniavov et al. (1970), which turns

---

10. For this reason Faye's studies (1964, 1972a, 1972b) are essential too, even in the narrative point of view of events.

11. The essential point of reference for all academics about the discursive *corpus* is Klemperer (1946), to whom the original acronym LTI (*Lingua Tertii Imperii)* was attributed; it was coined by the German philologist to be ironical about the frequent use of initials in the national socialist language.

out to be a landmark for the formalist analyses conducted on the speeches of the Russian revolutionary.

Thanks to quantitative methods, Cotteret & Moreau (1969) and Cotteret, Eméri, Gerstlé & Moreau (1976) draw up the last De Gaulle, Giscard d'Estaing and Mitterand's lexical repertoire. Following a contrastive approach, Roche (1979) analyses the semantic register of the last two presidents, which was adopted during the Elysium's French campaign in May 1974; instead, Singer (1976) takes into consideration the different linguistic uses of some White House's candidates.

Also, Noack (1970) analyses the typical features of the socio-political terminology of the subversive Jules Vallès; Cerroni (1978: 11–83) gathers 131 terms of Gramsci's vocabulary, with relative contextualizations, in order to understand the conceptual cruxes of both the philosopher and politician; Zolli (1980) conducts research about the elevated language of the first President of the Italian Republic, Luigi Einaudi, of which he provides a wide repertoire of lemmas; finally De Mattei (1976) gives attention to the language of the founder of the art of politics, par excellence, Niccolò Machiavelli, analysing in particular the lexemes' value which characterize *Il Principe* (*wisdom, knowledge, judiciousness* and *prudence*).

The single personalities' words contrast with those by collective subjects such as parties, movements, trends, groups, associations, etc.: all of them firmly codify their common elocutionary *habitus* which is recognized by citizens. Several studies focused their attention on this type of sign production, of which we will give just some examples. Bartholmes (1964) and Römer (1964) analyse some lexical peculiarities such as the isotopy of movement and the lexeme *Volk* which are resolutely present in the S.D.E. (*Sozialistiche Einheitspartei Deutschlands*) code from Eastern Germany; Bartholmes also (1970) traces the profiles of lemmas that are recurrent within the socialist terminology (*Bruder, Bürger, Freund, Genosse*); Jäger (1970) and von Weiss (1974) analyse vocabulary, slogans and the Left's watchwords.

Still talking about the semantic forms typical of specific parties, we mention Lantella's two studies (1973a, 1973b) on the presence of different lexemes such as *democrazia* and *libertà*, as well as collective terms, in the main documents with which the Italian parties have spread their programmes for the political elections on May 7, 1972; Labbé's works (1977, 1980) about the characteristics of the French Communist discourse, analysed according to the lexicographic criteria; Gerstlé's text (1979) on stereotypes and socialist language innovations.

As for movements and groups broadly speaking, Bartholomes (1966, 1968) and Orioles (1979) respectively analysed the lemma *Genosse* and its Italian *compagno* within the Italian and German labour movement's lexical code; Barat (1971) analyses the enemies' vocabulary structure of the Commune de Paris as well as the generative and paradigmatic process of the syntagma *La Commune de Paris*; Leclerc (1973) analyses the struggles among some political groups in

Indonesia and their linguistic consequences; Violi (1973) analyses the lexical strategies in *Lotta Continua*; finally Battistini (1976) studies the language of Bologna's Resistance conveyed through leaflets.

4.   The political language of the two Germanies

Starting from the Seventies, when the impossibility of an immediate return to the unification of Germany becomes even more evident, and concurrently the ideological and sociopolitical gap between the two States is more and more stressed, we can see, both in the Federal Republic of Germany and the Democratic one, a flourishing of studies pointing out the linguistic discrepancies, highliithing the different contextual uses, especially in the troubled field of civil life: the political involvement. These contrastive works reflect a particular postwar behaviour: the common desire of both countries to configure their own identity and consequently to consolidate the enemy's image.

Therefore two dictionaries were born (two different editions of *The Duden* for the West and the East); they aimed at justifying this need based on the metalinguistic codification of the respective allocations of meaning. Regarding this specific subject, it is really interesting to fully understand how these two contradictory realities reflect in the *langue*, thus we refer to Betz (1960); Schubert & Hellmann (1968), Jurgensen (1970) and Badura's (1971) contributions. The sociopolitical semantic fields that are more complex are analysed in the double Western and Marxist-Leninist sense, not before drawing up comparative topic lists.

The different specificities in the political lexicon adopted in the two Germanies, with respective connotations and different significational functions, have been the subject of elaborate research since 1961. By the way, one of the most symptomatic phenomena is the polysemous aspect of words. Indeed, different ideologies give different meanings to common terms: e.g. *Friedliche Koexistenz*, *Friedensicherung*, *Freiheit* and some hundreds of other symbolic words are interpreted in a totally different way both in the East and the West.

After selecting among academics who have worked on various linguistic problems in the two States, we refer to: Moser (1962), the first really systematic study about this topic; Ihlenburg (1964); Pelster's exemplary text (1966); Dieckmann (1967); Schippan (1967); Pizarczyk (1969) about the specific results of the mechanism of euphemism; Hellmann's reading (1973) which gathers a variety of interesting analysis about different semantic foundations of the political language in the two countries; finally Loewer (1975) that gives an overview of the main research about the different points of view.

The lexical choices and the official linguistic uses adopted in the Democratic German Republic in the different areas where the social and political involvement of the community takes place, are subject of the strict and accurate research by Reich (1968). Besides giving a rich and accurate glossary (Reich 1968: 11–231) of the language authorized by the ex-DDR, of which the academic analyses the Russian influences for each single item, forms and structures of the propagandistic vocabulary are analysed, as well as definitions, euphemisms and terms that disappear from the verbal political scenery.

Since the last two decades of the twentieth century to the present day the research on political discourse (not forgetting the lexical-semantic approach) will especially commit to the rhetoric, pragmatic, communicative processes, including the reflection on mass-media textual influences as well as current social networks.

References

Badura, Bernhard. 1971. *Sprachbarrieren. Zur Soziologie der Kommunikation*. Stuttgart: Frommann.

Barat, Michel. 1971. Le vocabulaire des ennemis de la Commune. Du 28 mai 1871 au mois de juin 1872 à travers les teste des comptes rendus des conseils de guerre. *La Pensée* 156: 52–67.

Barny, Robert. 1978. Mots et choses chez les hommes de la révolution française. *La Pensée* 202: 96–115.

Bartholmes, Herbert. 1964. *Das Wort "Volk" im Sprachgebrauch der SED*. Düsseldorf: Schwann.

Bartholmes, Herbert. 1966. Das Wort "Parteigenosse" im Sprachgebrauch der Arbeiterbewegung. *Muttersprache* 76: 148–153.

Bartholmes, Herbert. 1968. Das Wort "Genosse" samt seiner Zusammensetzungen im Sprachgebrauch der deutschen Arbeiterbewegung. *Muttersprache* 78: 193–222.

Bartholmes, Herbert. 1970. *Bruder, Bürger, Freund, Genosse und andere Wörter der sozialistischen Terminologie. Wortgeschichtliche Beiträge*. Göteborg: Acta Universitatis Gothoburgensis.

Battistini, Andrea. 1976. Lingua e oratoria nei volantini della Resistenza bolognese. In *L'Emilia Romagna nella guerra di liberazione, IV: Crisi della cultura e dialettica delle idee*. Atti del Convegno (Bologna, 2–5 aprile 1975), Annamaria Andreoli, (ed.), 331–364. Bari: De Donato.

Bergsdorf, Wolfgang. 1978. Die Sprache der Diktatur und ihre Wörter. Zur Technik nationalsozialistischer und kommunistischer Sprachlenkung. *Deutschland Archiv* 11(12): 1299–13l3.

Betz, Werner. 1960. Der zweigeteilte Duden. *Der Deutschunterricht* 12(5): 82–98.

Blanchard, Marc Eli. 1980. *Saint-Just & Cie. La Révolution et les mots*. Paris: Nizet.

Bork, Siegfried. 1970. *Missbrauch der Sprache. Tendenzen nationalsozialistischer Sprachregelung*. Bern: Francke.

Carofiglio, Vito. 1975. La "langue de la liberté". Notes sur le langage politique de Marat. In *Le langage politique de la Révolution française*, Vito Carofiglio (ed.), 5–18. Milano: CUESP.

Cerroni, Umberto. 1978. *Lessico gramsciano*. Roma: Editori Riuniti.

Cillán Apalategui, Antonio. 1970. *El léxico político de Franco en las Cortes Españolas*. Zaragoza: Imprenta Tipo-Línea.

Cobet, Christoph. 1973. *Der Wortschatz des Antisemitismus in der Bismarckzeit*. München: Fink.

Cofrancesco, Dino. 1978. Un'analisi del Dizionario di Politica. *Il Politico* 43(1): 106–138.

Cotteret, Jean-Marie & Moreau, René. 1969. *Le vocabulaire du Général De Gaulle. Analyse statistique des allocutions radiodiffusées 1958–1965*. Paris: Armand Colin.

Cotteret, Jean-Marie, Eméri, Claude, Gerstlé, Jacques & Moreau, René. 1976. *Giscard d'Estaing / Mitterand. 54 774 mots pour convaincre*. Paris: Presses Universitaires de France.

De Mattei, Rodolfo. 1976. Sapienza, sapere, saviezza e prudenza nel Machiavelli. *Lingua nostra* 37(1–2): 14–17.

De Mauro, Tullio. 1958. Il significato di 'democrazia'. *Il Ponte* 14: 40–47. Ristampa 1971. In *Senso e significato. Studi di semantica teorica e storica*, 228–238. Bari: Adriatica.

Desideri, Paola. 2009. Origini e sviluppi delle analisi e delle teorie sul linguaggio politico (1920–1960). *Linguistica* 49(2): 41–53. https://doi.org/10.4312/linguistica.49.1.41-53

Desideri, Paola & Marcarino, Aurelia. 1980. *Testualità e tipologia dei discorso politico. Bibliografia*. Roma: Bulzoni.

Dieckmann, Walther. 1964. *Information oder Überredung. Zum Wortgebrauch der politischen Werbung in Deutschland seit der Französischen Revolution*. Marburg: Elwert.

Dieckmann, Walther. 1967. Kritische Bemerkungen zum sprachlichen Ost-West-Problem. *Zeitschrift für deutsche Sprache* 23(3): 136–165.

Dieckmann, Walther. 1969. *Sprache in der Politik. Einführung in die Pragmatik und Semantik der politischen Sprache*. Heidelberg: Winter.

Drigeard, Gabrielle, Geffroy, Annie, Laplace, Roselyne, Tournier, Maurice & Wagner, Robert-Léon. 1977. *Vocabulaire socio-politigue de 1770 à 1814. Matériaux pour l'histoire du vocabulaire français*, 2ᵉ série, Vol. 11. Paris: Klincksieck.

Dubois, Jean. 1962. *Le vocabulaire politique et social en France de 1869 à 1872. A travers les œuvres des écrivains, les revues et les journaux*. Paris: Larousse.

Dubois, Jean. 1969. Lexicologie et analyse d'énoncé. *Cahiers de Lexicologie* 15: 115–126.

Enzi, Aldo. 1971. *Il lessico della violenza nella Germania nazista*. Bologna: Pàtron.

Faye, Jean Pierre. 1964. Langages Totalitaires: Fascistes et Nazis. *Cahiers Internationaux de Sociologie* 36: 75–100.

Faye, Jean Pierre. 1972a. *Langages totalitaires. Critique de la raison l'économie narrative*. Paris: Hermann.

Faye, Jean Pierre. 1972b. *Théorie du récit. Introduction aux 'langages totalitaires'*. Paris: Hermann.

Geffroy, Annie. 1973. L'étendue du vocabulaire chez Hébert et Robespierre. *Cahiers de Lexicologie* 22: 96–107.

Geffroy, Annie. 1974. Formes de base et forms spécifiques dans le discours robespierriste. *Cahiers de Lexicologie* 25: 96–136.

Geffroy, Annie. 1980. Trois successeurs de Marat pendant l'été 1793: Analyse lexicométrique des spécificités. *Mots* 1: 167–187. https://doi.org/10.3406/mots.1980.1009

Gerstlé, Jacques. 1979. *Le langage des Socialistes*. Paris: Stanké.

Glimm, Gesine. 1961. Die Kampfsprache Lenins. *Jahrbücher für Geschichte Osteuropas* 9: 511–540.

Guilhaumou, Jacques. 1975. 'Moment actuel' et processus discursifs. Le 'Père Duchesne' d'Hébert et le 'Publiciste de la République française' de Jacques Roux (14 juillet – 6 septembre 1793). *Bulletin du Centre d'Analyse du Discours de 1'Université de Lille III* 2: 147–173.

Guilhaumou, Jacques. 1980. La mise en scène de l'anglais dans le Père Duchesne d'Hébert (juillet 1793 – février 1794). Le jacobinisme à 1'épreuve du paradoxe. *Komparatistische Hefte* 2: 102–115.

Guilhaumou, Jacques, Maldidier, Denise, Prost, Antoine & Robin, Régine. 1974. Langages et idéologies. *Le Discours comme objet de l'Histoire*, Paris: Les Éditions Ouvrières.

Harnisch, Anna. 1974. Zur Analyse des Inhalts politischer Fachwörter. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 27(1–3): 65–72.

Hellmann, Manfred (ed.) 1973. *Zum öffentlichen Sprachgebrauch in der Bundesrepublik Deutschland und in der DDR. Methoden und Probleme seiner Erforschung*. Düsseldorf: Schwann.

Ihlenburg, Karl Heinz. 1964. Entwicklungstendenzen des Wortschatzes in beiden deutschen Staaten. *Weimarer Beiträge* 10(3): 374–397.

Jäger, Siegfried. 1970. Linke Wörter. Einige Bemerkungen zur Sprache der APO. *Muttersprache* 80: 85–105.

Jurgensen, Manfred. 1970. Die Sprache im zweigeteilten Duden. *Zeitschrift für deutsche Sprache* 26: 42–59.

Klemperer, Victor. 1946. *LTI. Die unbewältigte Sprache. Aus dem Notizbuch eines Philologen*. Darmstadt: Melzer.

Krauss, Werner. 1970a. Patriote, patriotique, patriotism am Ende des Ancien Régime. *Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg* 19(3–4): 79–83.

Krauss, Werner. 1970b. Zur Bedeutunggsentwicklung von "révolution". *Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg* 19(3–4): 87–90.

Labbé, Dominique. 1977. *Le discours comuniste*. Paris: Presses de la Fondation nationale des sciences politiques.

Labbé, Dominique. 1980. Le discours comuniste. *Revue Française de Science Politique* 30(1): 46–77. https://doi.org/10.3406/rfsp.1980.393878

Lange, Gerhard. 1968. Sprachform und Sprechform in Hitlers Reden. *Muttersprache* 78: 342–349.

Lantella, Lelio. 1973a. La lotta delle parole. Appunti su una campagna elettorale. *Il Politico* 38(1): 33–72.

Lantella, Lelio. 1973b. La lotta delle parole. Appunti su una campagna elettorale. Parte seconda: problemi di quantificazione. *Il Politico* 38(2): 244–279.

Lazzari, Giovanni. 1975. *Le parole del fascismo*. Roma: Argileto editori.

Leclerc, Jacques. 1973. Vocabulaire politique et répression politique: Un exemple indonésien. *Annales (E.S.C.)* 28(2): 407–428.

Leso, Erasmo. 1970. Il 'cittadino' nel triennio rivoluzionario (1796–1799). *Lingua nostra* 31(4): 111–117.

Leso, Erasmo. 1973. Aspetti della lingua del fascismo. Prime linee di una ricerca. In *Storia linguistica dell'Italia nel Novecento*, Atti del V Convegno Internazionale di Studi della Società di Linguistica Italiana (Roma 1–2 giugno 1971), Maurizio Gnerre, Mario Medici & Raffaele Simone (eds), 139–158. Roma: Bulzoni.

Leso, Erasmo. 1976. Lingua politica alla fine del Settecento: Storia di "moderato". *Lingua nostra* 37(1–2): 1–7.

Leso, Erasmo, Cortelazzo, Michele A., Paccagnella, Ivano & Foresti, Fabio. 1977. *La lingua italiana e il fascismo*, Introduzione di Luigi Rosiello. Bologna: Consorzio Provinciale Pubblica Lettura.

Loewer, Uwe. 1975. Die deutsche Sprache in der DDR. Ein Klotz am rechten Bein der BRD – Sprachwissenschaft. *Linguistische Berichte* 38: 81–105.

Mancuso, Josy. 1975. Les chansons révolutionnaires (1789–1794). *Bulletin du Centre d'Analyse du discours de l'Université de Lille III* 2: 87–102.

Medici, Mario. 1976. Sui francesismi politici e rivoluzionari del periodo giacobino in Italia. *Archivio Glottologico Italiano* 61: 193–199.

Moser, Hugo. 1962. *Die sprachlichen Folgen der politischen Teilung Deutschlands*. Düsseldorf: Schwann.

Mueller, Claus. 1973. The Politics of Communication. *A Study in the Political Sociology of Language, Socialization, and Legitimation*. Oxford: OUP.

Noack, Rudolf. 1970. Zum sozialen und politischen Wortschatz bei Jules Vallès. *Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg* 19(3–4): 209–214.

Orioles, Vincenzo. 1979. Il termine 'compagno' nella storia del movimento operaio. Gli anni della prima internazionale (1864–1876). *Officina dello storico* 1: 107–115.

Pelster, Theodor. 1966. *Die politische Rede im Westen und Osten Deutschlande. Vergleichende Stiluntersuchungen mit beigefügten Texten*. Düsseldorf: Schwann.

Pisarczyk, Karl. 1969. Der Euphemismus. Politische Sprache in DDR und BRD. *Krankfurter Hefte* 24(2): 107–116.

von Polenz, Peter. 1967. Sprachpurismus und Nationalsozialismus. In AA.VV., *Germanistik – eine deutsche Wissenschaft*, 111–160. Frankfurt: Suhrkamp.

von Proschwitz, Gunnar. 1966. Le vocabulaire politique au XVIII^c siècle avant et après la Révolution. Scission ou continuité? *Le Français Moderne* 34(2): 87–102.

Quasthoff, Uta. 1973. *Soziales Vorurteil und Kommunikation. Eine sprachwissenschaftliche Analyse des Stereotyps*. Frankfurt: Athenäum.

Rebollo Torío, Miguel Ángel. 1975. *El lenguaje de la derecha en la 2ª República*. Valencia: Fernando Torres.

Rebollo Torío, Miguel Ángel. 1978. *Lenguaje y política. Introducción al vocabulario político republicano y franquista 1931–1971*. Valencia: Fernando Torres.

Reich, Hans H. 1968. *Sprache und Politik. Untersuchungen zu Wortschatz und Wortwahl des offiziellen Sprachgebrauchs in der DDR*. München: Hueber.

Renaudie, Jean-Luc. 1975. Le "Courrier" de Gorsas (4 août 1792 – 11 août 1792). *Bulletin du Centre d'Analyse du Discours de l'Université de Lille III* 2: 103–115.

Robin, Régine (ed.). 1973. *Histoire et linguistigue*. Paris: Colin.

Roche, Jean. 1979. Deux candidats, deux styles: Valéry Giscard d'Estaing et François Mitterand. (élections présidentielles, mai 1974). *Travaux de Linguistique et de Literature* 17(1): 273–290.

Roegele, Otto B. (ed.). 1972. *Kleine Anatomie politischer Schlagworte*. Osnabrück: Fromm.

Römer, Ruth. 1964. Das Motiv der Bewegung in der Sprache der SED. *Wirkendes Wort* 14: 170–183.

Ruetsch, Pierre. 1975. Le "Courrier" de Gorsas (25 mai 1793 – 31 mai 1793). *Bulletin du Centre d'Analyse du Discours de l'Université de Lille III* 2: 117–145.

Schippan, Thea. 1967. Die beiden deutschen Staaten und die Sprache. *Der Deutschunterricht* 20(1): 8–18.

Schubert, Arne & Hellmann, Manfred W. 1968. Duden aus Leipzig und Mannheim. *Deutsche Studien* 23: 248–263.

Siccardo, Francesco. 1978. La lingua politica nell'"Ancien Régime" (1660–1789). Progetto lessicografico. *Annali della Facoltà di Scienze Politiche Università di Genova*, 4–5: 439–495.

Simonini, Augusto. 1978. *Il linguaggio di Mussolini*. Milano: Bompiani.

Singer, Aaron. 1976. *Campaign Speeches of American Presidential Candidates*. New York NY: Ungar.

Šklovskij, Viktor, Tynjanov, Jurij N. et al. 1970. *Sprache und Stil Lenins*. Berlin: Volk und Welt.

Slakta, Denis. 1971. Esquisse d'une théorie lexico-sémantigue pour une analyse d'un texte politique (Cahiers de doléances). *Langages* 23: 87–134.

Vardar, Berke. 1973. *Structure fondamentale du vocabulaire social et politique en France, de 1815 à 1830*. Instanbul: Imprimerie de la Faculté des Lettres.

Violi, Patrizia. 1973. Il linguaggio di Lotta Continua. *Versus* 6: 105–126.

von Weiss, Andreas. 1974. *Schlagwörter der Neuen Linken. Die Agitation der Sozialrevolutionäre*. München: Olzog.

Winckler, Lutz. 1970. *Studie zur gesellschaftlichen Funktion faschistischer Sprache*, Frankfurt: Suhrkamp.

Zimmermann, Hans Dieter. 1969. *Die politische Rede. Zum Sprachgebrauch Bonner Politiker*. Stuttgart: Kohlhammer.

Zolli, Paolo. 1964. Retrodatazioni di francesismi settecenteschi. *Lingua Nostra* 25(1): 11–17.

Zolli, Paolo. 1965. Francesismi nel linguaggio politico italiano alla fine del '700. *Lingua Nostra* 26(1): 16–19.

Zolli, Paolo. 1980. Luigi Einaudi e la lingua italiana. *Lingua Nostra* 41(2–3): 89–99.

CHAPTER 24

Dialects and neuroscience

A first critical review

Mirko Grimaldi

University of Salento

Dialectal variation has been constanly at the core of linguistic investigation. On the base of the few neurophysiological studies that explored dialectal data, this article examines if and how dialectal variation may contribute to the advancing of the neurobiology of language. Evidence suggest that the neural investigation of allophonic variation generated by phonological rules is very useful for the progress of the neurobiology of language. Also, all kinds of parametric variations characterizing dialects are well suited to this aim. Finally, I outline some general and detailed questions that need to be addressed making use of a new methodology: the oscillatory rhythms approach.

Keywords: dialectal variation, neuroscience, ERPs, oscillatory rhythms, neurobiology of language

1.   Introduction

At the end of the nineteenth century, Neogrammarian, in searching for the general principles of sound changes, recognized that dialects evidence would have been relevant. Following the hypothesis that 'sound changes are exceptionless', data on dialect variation might have proved that the evolution of ancient languages into modern languages had followed regular phonological changes. This line of research led to the development of *dialect geography* (or more simply *dialectology*) and the realization of linguistic atlas.[1]

First findings of dialectal investigations, however, seemed to undermine any claim on the exceptionlessness of sound changes. Indeed, they revealed a variation that was unimaginable beforehand and this heterogeneity seemed almost random to the first researchers: speakers of neighboring localities gave wildly inconsistent

---

[1] See a collection of linguistic atlas here: <http://www.accademiadellacrusca.it/en/link-utili/collections-linguistic-atlas>

responses to questioned items, and sometimes even they were inconsistent in their own responses from day to day. Only in the second half of the twentieth century such variability has become the focus of linguistic theorizing, with its own principles and rule-governed systematicity (Chamber & Trudgil 2004; Kenstowicz 1994; Chambers & Schilling 2013). Also, contemporary theoretical linguistics continue to look at dialect variation as a crucial source of data useful to verify, refine, and develop theoretical models: among many others, representative examples of this approach are the *Atlante Sintattico d'Italia* (Poletto 2000), the morpho-syntactic investigation of Romance and Romansh dialects (Manzini & Savoia 2005, 2007), and the *Cartographic enterprise* (Cinque & Rizzi 2008).

At the end of the twentieth century another discipline, the cognitive neuroscience, in searching for the neural correlates of human cognitive abilities, was undertaken with a *cartographic enterprise*. Neuroscientists were initially committed to develop useful neural maps, a sort of neural atlas of cognitive functions. The idea was that localization on a map of brain activity amounts to explanation functionality of the brain. It appeared that the localization of psychological functions had become the epistemological driving force. So, brain images were assumed to constitute a privileged type of evidence that form the basis for the explanation of mental life (Poeppel 2008). In this enterprise, a classical approach was to map, for example, the neural structures involved during speech perception and production, during the elaboration of words (semantic categorization and decision), or during sentences comprehension, etc.

However, the conclusions deduced from these studies are inherently *localizationistic* in nature. In other words, they describe linguistic cognitive functions as being localized in focal brain regions. Furthermore, it is tacitly assumed that categories, notions and principles developed within theoretical linguistics can be validated through neurophysiological and neuroimaging methods: i.e., that ontological entities such as 'phoneme', 'morpheme', 'word', 'syntax', etc., can be neurobiologically measured. Clear limits of such an approach are the lack of an explanatory theory that permits to correlate linguistic ontologies with neuronal ontologies: the risk is to fall into a sterile reductionism (Poeppel & Embick 2005; Grimaldi 2012). Thus, a direct reduction of the linguistic primitives (phonemes, syllables, words, etc.) into the activation of neural areas, is a limitation in the progress of an integrated study of language and brain. Indeed, to account for the combinatorial flexibility required for production and perception language processes, we need to think about linguistic computational and representational processes as consisted in dynamically associated assemblies of neurons rather than individual specialized cells. Neurons are grouped into assemblies, and their responses are connected through a binding mechanism based on the transient and precise synchronization of their discharges and active brain states.

Thus, researchers in the field of cognitive neuroscience of language are become aware that we need to proceed in developing an inter-field theory that integrates and bridges fields rather than establishing one complete, unified theory. Inter-field theories can be generated when two fields share an interest in explaining different aspects of the same phenomenon to build solid knowledge and relations between the fields. This perspective advocates integration rather than reduction: according-ly, an inter-field theory should interconnect well-established linguistic computa-tional primitives with neurophysiological computations responsible for represen-tational processes at the light of the knowledge reached within each research area. This demanding task will lead us to progressively create epistemological bridges between different disciplines. More precisely, the task ahead is to characterize this kind of linked computations and find out how they work in concert producing lin-guistic behaviors: and step by step it is probable that the 'neurobiology of language' may stand on its own feet integrating the two research traditions and producing an inter-theoretic framework (Grimaldi in press a).

As it happens for Neogrammarian, in searching for the neural principles governing language processing, also neuroscientists are beginning to investigate dialect variation. So, it seems that dialects represent a sort of test bed or of addi-tional resources to test theoretical assumptions and verify empirical data. Indeed, abstract linguistic structures (phonemes, words and their morphosyntactic rela-tions) are continuously subjected to systematic variation within linguistic systems: for example, phonemes may be deleted or assimilated in specific contexts to other phonemes, and so on for words formations and syntactic structures. Accordingly, dialect data may reveal the deep nature of neural computations intrinsic to natural languages.

With this work, I aim to provide a first critical review of the few studies ad-dressed dialectal variation from a neurobiological perspective. We will try to un-derstand what dialects have to say on the neurobiology of language and if and how their investigation may be useful for the advance of research in this field.

2. Investigating the auditory brain: Techniques and methods

Two techniques are widely used to study the time course and neurophysiological underpinnings of language processing: electroencephalography (EEG) and mag-netoencephalography (MEG), as they are the most powerful non-invasive tools with high temporal reliability (Roberts et al. 2000). Recently, also electrocorticog-raphy (ECoG) – an invasive approach used in clinical contexts where pre-surgical evaluation of cognitive processes is needed – is increasingly used to directly record auditory activity (Leonard & Chang 2016).

MEG and EEG research into language processing is based on event-related potentials (ERPs) and event-related magnetic fields (ERMFs) recorded while the subjects are performing a cognitive task with sensor collocated on the scalp (see below). Thereby stimuli are presented to subjects and markers are set into the EEG trace whenever a stimulus is presented. Then a short epoch of EEG/MEG around each marker is used to average all these segments. This is based on the logic that in each trial there is a systematic brain response to a stimulus. Practically, this means that one typically repeats a given experimental paradigm a number of times (say, > 30 times), and then one averages the EEG/MEG recordings time-locked to the experimental event.

However, this systematic response cannot be seen in the raw EEG, as there it is overlaid by unsystematic background activity (which is generally considered as noise). By averaging all the single epochs that are time-locked to the experimental event, only the systematic brain response should remain (i.e., those generate neural action potentials related to the stimuli), but the background EEG/MEG should approach zero (Sauseng & Klimesch 2008). The noise (which is assumed to be randomly distributed across trials) diminishes each time a trial is added to the average, while the signal (which is assumed to be stationary across trials), gradually emerges out of the noise as more trials are added to the average. These brain responses, named ERPs and ERMFs reflect the summated activity of network ensembles active during the task. ERPs/ERMFs are characterized by specific patterns called 'waveforms' (or 'components'), which are elicited around 50–1000 ms starting from the onset of the stimulus and show positive (P) and negative (N) oscillatory amplitudes (i.e., voltage deflections). For instance, P100, N100, P200, P300, N400, P600 (or P1, N1, P2, and so on) are the principal components elicited during language processing starting from sound perception to semantic and syntactic operations. So, this techniques provide millisecond-by-millisecond indices of brain functions and therefore provide excellent temporal resolution (Luck 2005).

For the goal of the present paper, it is necessary to focus on the N1 component, with its magnetic counterpart N1m, and mismatch negativity (MMN), with its magnetic counterpart MMNm. N1/N1m is a negative peak between 70 and 150 ms after the onset of an auditory stimulus (cf. Figure 1) that appears to be involved in the basic processing of speech sounds in auditory cortices. It seems that the amplitudes and the latencies of the N1/N1m are relevant markers reflecting the cortical encoding of acoustic features of incoming speech sounds. The source location of the N1/N1m responses along the auditory planes seems to be driven by the spectral properties that are linguistically salient, e.g., the F1/F2 ratio for vowels, or the place of articulation for consonants, and then it may represent a good tool to investigate auditory cognitive processes (Manca & Grimaldi 2016).
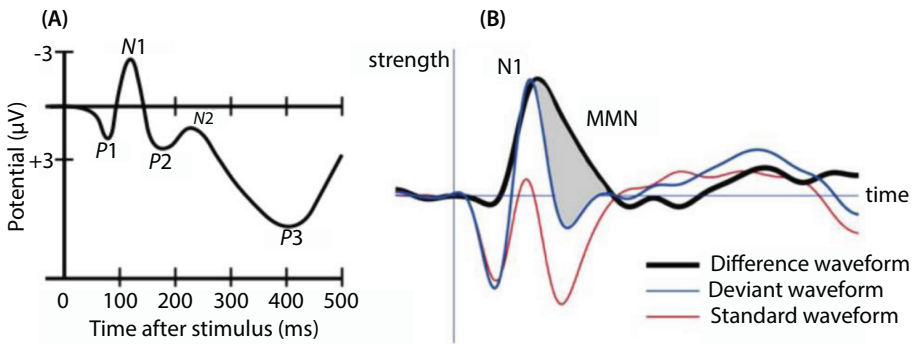
Figure 1. (A) Representation of the auditory N1 wave evoked from EEG to an auditory stimulus. The peak around 100 ms post-stimulus onset, measured in microvolts (μV) is evidenced (adapted from Lageman et al. 2012). (B) ERP waveforms evoked at a frontal scalp location by the standard and deviant sounds superimposed on the difference waveform in which the ERP to the standard has been subtracted from that to the deviant. The MMN appears as an enlarged negativity to the deviant sound as compared with the standard sound, following the N1 peak. Adapted from (Brattico 2006).

MMN/MNNm is a component temporally subsequent to the N1/N1m (cf. Figure 1), automatically and preattentively elicited by an acoustic change or by a rule violation between 150 and 250 ms post-stimulus onset (Näätänen 2001). Contrary to the N1/N1m, it is generated in a passive oddball paradigm, where subjects listen to frequent (standard) stimuli interspersed with infrequent (deviant) stimuli and attend to a secondary task (e.g., watching a silent movie). MMN/MMNm is visible by subtracting standard responses from deviant responses to acoustic stimuli and its amplitude seems to be directly correlated with the discriminability of the two stimuli involving both acoustic change-detection processes and phoneme-specific processes. One interpretation of the MMN is such that the standard activates a memory trace with which the deviant is compared. The more variability that occurs within the standards, the more likely that listeners construct an abstraction from the actual acoustic stimulus. Thus, variable standard tokens are commonly used to engage phonetic and phonological effects that involve more than just acoustic stimulus information. Furthermore, it seems that the MMN response is modulated by top-down information pertaining to language-specific long-term representations of speech sounds.

This component has been exploited to investigate (i) the categorical representation of phonemes in the subjects' mother tongue; (ii) if the acoustic signal is mapped onto lexical representations through different levels of featural representation; in this case, N1m and MMNm have also been used together, and (iii) if phonemic representations may eventually develop during second language acquisition.

3. Dialect in the brain

3.1 Neurophysiology of tonal dialects

A first, tentative, line of research seems interested to understand how tonal dialects are processed by the brain, since pitch contours may reflect lexical contrast. Fournier et al. (2010) investigated native speakers of the Roermond tonal dialect (in the Netherlands), compared with a control group of speakers of Standard Dutch (a non-tone language). They used three conditions: (i) lexical, where the pitch contour differences between standard and deviant stimuli reflected differences between lexical meanings; (ii) intonational, where the stimuli differed in their discourse meaning; (iii) combined, where the stimuli differed both in their lexical and discourse meaning. The native Roermond listeners clearly exhibit a stronger MMNm response over the left temporal cortex for lexical contrasts, but a predominantly right response for the intonational contrasts. Conversely, Standard Dutch listeners showed a stronger MMNm over the left temporal cortex in all conditions. Accordingly, these data suggest that prosody patterns used to control lexical contrasts are processed in the right hemisphere, while when they are used to disambiguate intonational structures are processed in left hemisphere.

These findings were refined by Sato et al. (2013), who studied native speakers of standard (Tokyo) Japanese, which has a lexical pitch accent system, and native speakers of accentless dialects, which do not have any lexical tonal phenomena. However, accentless dialects speakers were exposed to the standard dialect since a young age, so they showed sensitivity and knowledge to lexical pitch accent. Also in this case, only the standard Japanese speakers showed left-dominant responses in temporal regions to pitch pattern changes within words. Thus, it seems that listener's language experience influences functional lateralization, even within dialects of the same language, since speakers of the standard Japanese dialect have a different perceptual pattern for sound contrasts compared to speakers of accentless dialects.

3.2 Auditory processing and dialect variation

A second considerable way to investigate the neural correlates of dialectal variation is looking at how auditory processes are affected by phonological features of dialects in respect of standard or other regional varieties.

For instance, standard American English speakers were exposed to variable tokens of the word *Hello*. A set of tokens were pronounced in the standard American English variety, while another set in the African-American English variety (Scharinger, Monahan & Idsardi 2011). The main difference between the

tokens consists in the acoustic properties of the vowel [ɛ], which in Standard American English is produced slightly more front in the mouth than its African-American English counterpart. The results showed specific MMNms responses only for the Standard American English 'Hello' tokens. Accordingly, the authors suggest that cues for dialectal background information are rapidly extracted from spoken words and processed in the same way as speaker voice information, since the attribution of a speaker to a dialectal community is necessarily mediated by speaker identity. Bottom-up acoustic information is then integrated with top-down knowledge of speaker and dialect categories in long-term memory.

Southern American dialects show a well-known phenomenon called *pin/pen* merger, so that /ɪ/ ([pɪn]) and /ɛ/ ([pɛn]) before nasal consonants are merged to [ɪ]. How merged dialect speakers process word contrast of non-merged dialects? American speakers with and without vowels merger were asked to discriminate perceptually between minimal pairs of words that contrasted in the critical vowel merger and minimal pairs of control words (Conrey, Potts & Niedzielski 2005). Compared with unmerged dialect speakers, merged dialect speakers were less able to make behavioral discriminations. Neurally, merged dialect speakers exhibited a reduced late ERP answer, named late positive component, to *pin/pen* stimuli, whereas the non-merged dialect speakers elicited a clear late positive component. This kind of neural answer suggests that unmerged dialect speakers were able to make a conscious decision on the word contrast while the merged speakers not.

Along this line, Brunellière et al. (2009) compared the perception of the French vowel contrast /ɛ/–/e/, which is no longer contrastive in the South of France but is in standard (northern) French, with a control contrast /ø/–/y/, which is present in both varieties. Standard French listeners were found to be slower and less accurate with the merging /ɛ/–/e/ contrast than the control one, as resulted by MMNms amplitudes. Further comparisons showed that the control vowels elicited differences across P200 (190–230ms), MMN (270–310ms), and a later time-period (372–486 ms), whilst for the merging vowels there was only a difference in the MMN. It was suggested that this was due to the difficulty in distinguishing the merging vowels at a phonemic processing level due to participants' exposure to merging accents, leading to a change in the organization of regions that selectively answer to native language vowels. This was supported by a follow-up study, where Brunellière, Dufour & Nguyen (2011) found that the /ɛ/–/e/ contrast induced different cortical topographies only in the standard French population, but not in the Southern French population. Correspondingly, by examining two French regional dialects, Dufour, Brunellière & Nguyen (2013) demonstrated that speech perception is highly linked to the listener's linguistic competence even within dialects of a single language.

A definitive confirmation of these findings was found by Lanwermeyer et al. (2016) who studied the /ȏa/-/ȏʊ/ contrast, which is a stable contrast in the Central Bavarian dialect, while, due to a merger phenomenon, only /ȏa/ occurs before obstruents in the neighboring Bavarian-Alemannic transition. Results showed that for Bavarian-Alemannic listeners the /ȏa/-/ȏʊ/ contrast evoke enhanced neural costs during sentence comprehension. These findings demonstrate a conscious mismatch detection and evaluation process with regard to the non-native lexemes. Thus, it seems likely that the usage of these non-native contrast causes problems during cross-dialectal communication, leading to negative feedback and to subsequently competence modifications.

If phonological dialect features drive the way in which other phonological properties are auditorily processes, it seems that different regional dialects pronunciations do not affect the access to semantic differences in preliterate children speaking one of two German language varieties (i.e., standard German and Swiss German), where one language variety corresponds more strongly to the German written norm than the other (Bühler et al. 2017a).

3.3   Allophonic variation within dialects

All these studies concentrate their attention on how listeners discriminate unfamiliar allophones: that is, allophones not present in their phonological system. An interesting question is whether allophones generated by phonological rules within a linguistic system are processed as unfamiliar allophones or whether they are perceptually treated as phonemes inasmuch they are part of the speakers' grammar.

A first study showed that Koreans, who have the allophone [d], which occur only between voiced sounds, do not discriminated the allophone at the same way they do for the phoneme [t] which occur elsewhere (Kazanina et al. 2006). Conversely, Hacquard et al. (2007) observed that Spanish and French listeners MMNm response to the vowel pair [ɛ-e] did not differed significantly, despite the fact that the vowel pair is phonemic in French but allophonic in Spanish. Thus, these findings suggest that vocalic and consonantal allophones might be neurally processed in different way.

However, Miglietta et al. (2013) found that phonemes and allophones display the same discriminative patterns. In particular, they investigated an allophonic variant generated by a phonological process characterizing southern Salento dialects where the stressed low-mid front vowel [ɛ] is raised to its high-mid counterpart [e] when followed by the unstressed high vowel [i]. The same MMN amplitudes were elicited for both the allophonic and phonemic conditions, but a shorter latency was observed for the phonemic vowel pair suggesting a rapid access to contrastive sound properties in the phonological mode.

A confirmation of these data comes from the study of Bühler et al. (2017b). They studied native speakers of Standard German living in Switzerland and native speakers of Swiss German. The former speakers show a phonological process for which unvoiced stops undergo a categorical allophonic rule of aspiration before stressed vowels which turns them into aspirated [pʰ], [tʰ], [kʰ], while the later exhibit a rule for which [p], [t], [k] are lengthened in [pː], [tː], [kː] within inter-vocalic position. Findings highlighted that both phonemes and allophones were clearly discriminated by both groups of speakers as indexed by MMN amplitudes. Overall, the discrimination of allophones generated by specific rules of the grammar indicates that they are part of the knowledge of speakers and then of their memory representations. Therefore, more studied are need to clarify this interesting issue which has important theoretical implications as well.

4.   Discussion and further remarks

What do dialects have to say on the neurobiology of language? Until now, very few studies have investigated the dialectal variation from the neurobiological perspective. All these studies have addressed just the phonetic-phonological level of analysis. As far as I know, no studies have been published addressing other levels of linguistic variation (but see Bambini et al. 2017).

It seems that most of the studies reviewed (Sections 2.1, 3.2) have followed a model of investigation traced by the seminal work of Näätänen (1997). In that work, Finnish subjects were presented with the Finnish vowel [e] as the standard stimulus, while two other Finnish vowels [ö], [o] and the Estonian vowel [õ] were the deviant stimuli. It was found that the MMNm elicited by the Estonian [õ] was smaller in amplitude than that elicited by Finnish vowels, implicating the presence of permanent phonemic traces in the auditory cortex. These findings opened a new way for the investigation of phonemes-specific processes implicated both in first and second language acquisition (cf. Manca & Grimaldi 2016 for a review). However, when this line of research is applied to study dialect variation it is questionable whether it may contribute to the progress of the neurobiology of language. For instance, comparing the auditory processes of merged dialect speakers with auditory processes of unmerged dialect speakers is the same as to compare Finnish speakers with Estonian speakers. In this way, we may reach further proofs on the neural correlates of auditory categorical processing, but little advance in the comprehension of the functional neuronal network controlling language may be done.

From this perspective, it seems more useful to investigate allophonic variation (cf. Section 2.3). The allophonic variation generated by phonological processes is at the core of the development of phonology as a branch of linguistics distinct

from phonetics (Kenstowicz 1994). So, it represents a well-established linguistic computational primitive that may be interconnected with neurophsysiological computations. Speakers are unconscious of the systematic rules that modify a segment depending on the context in which it finds itself generating a conditioned allophone: this unconsciousness, however, does not imply that allophones are not involved in learning and memory processing. In fact, if the perceptual representations computed from speech and encoded in memory representations are only phonemic, language-specific conditioned allophones could not be acquired in first or second language acquisition because of language-specific phonological rules and not universal co-articulation adjustments (Calabrese 2102). To acquire the Florentine variety, for example, a newborn must learn the rule for which voiceless stops /k t p/ are subjected to intervocalic spirantization ($\rightarrow$ [x/h, θ, ɸ]), and as consequence he/she must develop a memory representation of the allophonic variants.

What is inherent in dialects (and of course in languages) is the notion of *variation*; this variation is only apparently puzzled, since it is governed by rules. How the brain controls the acquisition of these rules? How these rules are instantiated in the neuronal network that controls language competence? This perspective has, of course, to be extended to other levels of languages. Until now, the neural investigation of linguistic computations and representations has been dominated by general questions as, for instance: what are the neural correlates of phonemes, words, morphemes, syntax? And so on. At a finer level of analysis, one can ask what is the neuronal reality of notions as *distinctive features* (Manca & Grimadi 2016), *recursion* (Pallier et al. 2010), *merge* (Rizzi 2012; Nelson et al. 2017), etc. This kind of findings are surely useful for the advancing of the neurobiology of language, but the time is ripe that researchers begin to investigate in parallel the neuronal reality of dialectal variation. This way may offer a unique opportunity to investigate the neural basis of linguistic competence and of specific parameters that generate linguistic variation.

A general question to address may be the following:

Does linguistic variation involve neuronal processes distinct from other basic forms of linguistic computations and representations?

Starting from this general issue, other questions arise:

How parameters that control linguistic variation are acquired, fixed, and, then, computed by the brain?
Are there neurophysiological patterns specifically entrusted with allophonic, morphological, syntactic, and other kind of linguistic variation?

> How phonological, morphological, syntactic, and semantic parameters gener-
> ated by grammatical rules are functionally computed and represented by the
> neuronal network designed to control a grammar, and how are they related
> with other cognitive functions?

These fundamental issues may be today investigated thanks to a new perspective opened within the EEG approach: that is, the event-related oscillations, connected with classical delta (~ 0.5–4 Hz), theta (~ 4–10 Hz), alpha (~ 8–12 Hz), beta (~ 12–30 Hz), and gamma (~ 30–100 Hz) rhythms. Event related oscillations reflect the extent to which the underlying neuronal activity synchronizes. Synchronization and de-synchronization are related to the coupling and uncoupling of functional networks in cortical and subcortical areas of the brain (see e.g., Varela et al. 2001). This aspect, of course, is related to how different types of information, which are stored in different parts of the network, are integrated during computational and representational processes. Importantly, elements pertaining to the same functional network are identifiable as such they fire synchronously at a given frequency. This frequency specificity allows the same neuron (or neuronal pool) to participate at different times in different representations. Hence, synchronous oscillations in a wide range of frequencies are considered to play a crucial role in dynamical linking areas that are part of the same functional network. Crucially, in addition to recruiting all the relevant network elements, oscillatory neuronal synchrony serves to bind together the information necessarily represented online during the elaboration of different elements.

With regard to the perspective of investigation here pursued, we may hypothesize that the same oscillatory rhythms control both basic linguistic operations and parametric variation, thanks to synchronization and de-synchronization processes dynamically binding different information within a neuronal network. Actually, there are solid evidence that oscillatory rhythms do not have single functions (Başar et al. 2001): it is thus unlikely that, each rhythm have a single role in language processing. For instance, gamma bands have been found involved in sentence level semantic unification, but also in speech perception and production (Grimaldi in press a, b).

In conclusion, it seems that the investigation of dialectal variation represents an inescapable point of reference when progresses have to be reached both from a theoretical and empirical perspectives. Addressing the investigation of dialectal variation from a neurocognitive perspective can be very productive to coherently integrate theoretical linguistics within the neurobiology of language.

Acknowledgment

References

Bambini, Valentina, Canal, Paolo, Breimaier, Federica, Meo, Domenico, Pescarini, Diego & Loporcaro, Michele 2017. How ERPs can reflect morphosyntactic change: A study on gender fading in a southern Italo-Romance dialect. Poster presented at the 30th CUNY conference, Boston.

Başar, Erol, Başar-Erogluc, Canan, Karakaş, Sirel & Schürmann, Martin 2001. Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology* 39: 241–248. https://doi.org/10.1016/S0167-8760(00)00145-8

Brattico, Elvira. 2006. *Cortical Processing of Musical Pitch as Reflected by Behavioural and Electrophysiological Evidence*. Helsinki: Yliopistopaino.

Brunellière, Angèle, Dufour, Sophie, Nguyen, Noël & Frauenfelder, Ulrich Hans. 2009. Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition* 111: 390–396. https://doi.org/10.1016/j.cognition.2009.02.013

Brunellière, Angèle, Dufour, Sophie & Nguyen, Noël. 2011. Regional differences in the listener's phonemic inventory affect semantic processing: a mismatch negativity (MMN) study. *Brain & Language* 117: 45–51. https://doi.org/10.1016/j.bandl.2010.12.004

Bühler, Jessica C., Schmid, Stephan & Maurer, Urs. 2017a. Influence of dialect use on speech perception: A mismatch negativity study. *Language, Cognition and Neuroscience* 32(6): 757–775. https://doi.org/10.1080/23273798.2016.1272704

Bühler, Jessica C., Waßmann, Franziska, Buser, Daniela, Zumberi, Flutra, Maurer. 2017b. Neural processes associated with vocabulary and vowel-length differences in a dialect: An ERP study in pre-literate children. *Brain Topography* 30: 610–628
https://doi.org/10.1007/s10548-017-0562-2

Calabrese, Andrea 2012. Auditory representations and phonological illusions: A linguist's perspective on the neurophysiological bases of speech perception. *Journal of Neurolinguistics* 25(5): 355–381. https://doi.org/10.1016/j.jneuroling.2011.03.005

Chamber, Jack K. & Trudgill, Peter. 2004. *Dialectology*. Cambridge: CUP.

Cinque, Guglielmo & Rizzi, Luigi. 2008. The cartography of syntactic structures. *CISCL Working Papers on Language and Cognition* 2: 43–59.

Conrey, B., Potts, P., Niedzielski, N. A. 2005. Effects of dialect on merger perception: ERP and behavioral correlates. *Brain and Language 95*: 435–449.

Dufour, Sophie, Brunellière, Angèle & Nguyen, Noël. 2013. To what extent do we hear phonemic contrasts in a non-native regional variety? Tracking the dynamics of perceptual processing with EEG. *Journal of Psycholinguistic Research* 42: 161–173
https://doi.org/10.1007/s10936-012-9212-8

Fourniera, Rachel, Gussenhoven, Carl, Jensen, Ohle & Hagoort, Peter. 2010. Lateralization of tonal and intonational pitch processing: An MEG study. *Brain Research* 1328: 79–88.
https://doi.org/10.1016/j.brainres.2010.02.053

Grimaldi, Mirko 2012. Toward a neural theory of language: Old issues and news perspectives. *Journal of Neurolinguistics* 24(5): 304–327. https://doi.org/10.1016/j.jneuroling.2011.12.002

Grimaldi, Mirko. To In press. From brain noise to syntactic structures: A formal proposal within the oscillatory rhythms perspective. *Linguistic Variation: Structure and Interpretation – A Festschrift in Honor of M. Rita Manzini in Occasion of her 60th Birthday*, Ludovico Franco & Paolo Lorusso (eds). Berlin: Mouton de Gruyter.

Grimaldi, Mirko. In press. The phonetics-phonology relationship in the neurobiology of language. *In From Sounds to Structures: Beyond the Veil of Maya*, Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds). Berlin: Mouton De Gruyter.

Hacquard, Valentine, Walter, Mary A. & Marantz, Alec. 2007. The effects of inventory on vowel perception in French and Spanish: An MEG study. *Brain & Language* 100: 295–300. https://doi.org/10.1016/j.bandl.2006.04.009

Jack, Chambers & Schilling, Natalie (eds). 2013. *The Handbook of Language Variation and Change*. Oxford: Wiley-Blackwell.

Kazanina, Nina, Phillips, Colin & Idsardi, Willima J. 2006. The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America* 103(30): 1138–1186. https://doi.org/10.1073/pnas.0604821103

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Lagemann, Lothar, Okamoto, Hidehiko, Teismann, Henning & Pantev, Christo. 2012. Involuntary monitoring of sound signals in noise is reflected in the human auditory evoked N1m response. *Plos One* 7(2): e31634. doi: https://doi.org/10.1371/journal.pone.0031634

Leonard, Matthew K., Chang & Edward, F. 2016. Direct cortical neurophysiology of speech perception. In *Neurobiology of Language*, Gregory Hickok & Steven L. Small (eds), 479–489, Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00039-0

Lanwermeyer, M. Henrich, K. Rocholl, M. J., Schnell, H. T., Werth, A. Herrgen, J. & Schmidt, J. E. 2016. Dialect Variation Influences the Phonological and Lexical-Semantic Word Processing in Sentences. Electrophysiological Evidence from a Cross-Dialectal Comprehension. *Frontiers. Psychology* 7:739. doi: https://doi.org/10.3389/fpsyg.2016.00739.

Luck, Steven, J. 2005. *An Introduction to the Event-related Potential Technique*. Cambridge MA: The MIT press.

Manca, Anna Dora & Grimaldi, Mirko. 2016. Vowels and consonants in the Brain: Evidence from magnetoencephalographic studies on the N1m in normal-hearing listeners. *Frontiers in Psychology* 7: 1413. https://doi.org/10.3389/fpsyg.2016.01413

Manzini, M. Rita & Savoia, Leonardo, M. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*. Torino: Dell'Orso.

Manzini, M. Rita & Savoia, Leonardo, M 2007. *A Unification of Morphology and Syntax. Investigations into Romance and Albanian dialects*. London: Routledge.

Miglietta, Sandra, Grimaldi, Mirko & Calabrese, Andrea. 2013. Conditioned allophony in speech perception: An ERP study. *Brain & Language* 126(3): 285–290 https://doi.org/10.1016/j.bandl.2013.06.001

Näätänen, Risto 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38: 1–21. https://doi.org/10.1111/1469-8986.3810001

Näätänen, Risto, Lehtokoski Anne, Lennes Mietta, Cheour Marie, Huotilainen Minna, Livonen et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385: 432–434. https://doi.org/10.1038/385432a0

Nelson, Matthew J., Karoui, Imen El, Giber, Kristof, Yang, Xiaofang, Cohen, Laurent, Koopman, Hilda, Cash, Sydney, et al. 2017. Neurophysiological dynamics of phrase-structure building during sentence processing. *PNAS* 114(18): E3669–E3678. https://doi.org/10.1073/pnas.1701590114

Pallier, Christophe, Devauchellea, Anne-Dominique & Dehaene, Stanislas. 2010. Cortical representation of the constituent structure of sentences. *PNAS* 108(6): 2522–2527. https://doi.org/10.1073/pnas.1018711108

Poeppel, David & Embick, David 2005. Defining the relation between linguistics and neuroscience. In *Twenty-First Century Psycholinguistics: Four Cornerstones*, Anne Cutler (ed.), 103–118. Mahwah NJ: Lawrence Erlbaum Associates.

Poeppel, David. 2008. The cartographic imperative: Confusing localization and explanation in human brain mapping. *Bildwelten* 6(1): 19–29.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. Oxford: OUP.

Rizzi, Luigi 2012. Core linguistic computations: How are they expressed in the mind/brain? *Journal of Neurolinguistics* 25: 489–499. https://doi.org/10.1016/j.jneuroling.2009.12.005

Roberts, Timothy P., Ferrari, Paul, Stufflebeam, Steven M. & Poeppel, David. 2000. Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology* 17: 114–129. https://doi.org/10.1097/00004691-200003000-00002

Sato, Yutaka, Utsugi, Akira, Yamane, Naoto, Koizumi, Masatoshi & Mazuka, Reiko. 2013. Dialectal differences in hemispheric specialization for Japanese lexical pitch accent. *Brain & Language* 127: 475–483 https://doi.org/10.1016/j.bandl.2013.09.008

Sauseng, Paul & Klimesch, Wolfgang 2008. What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioral Reviews* 32: 334–348. https://doi.org/10.1016/j.neubiorev.2008.03.014

Scharinger, Mathias, Monahan, Philip J., Idsardi, William J. 2011. You had me at "Hello": Rapid extraction of dialect information from spoken words. *NeuroImage* 56(4): 2329–2338. https://doi.org/10.1016/j.neuroimage.2011.04.007

Varela, Francisco, Lachaux, Jean-Philippe, Rodriguez, Eugenio & Martinerie, Jacques 2001. The brain web: phase synchronization and large-scale integration. *Nature Review Neuroscience* 2: 229–239. https://doi.org/10.1038/35067550

CHAPTER 25

Remarks on the vulnerability of grammar

Giovanna Marotta

University of Pisa

This paper deals with the special status of morphological features in a series of contexts that are crucial for the understanding of the architecture of human grammar: language pathology and impairment, L1/L2 acquisition, language evolution, language contact. The critical review of data referring to different domains and to different languages enlightens the special vulnerability of morphology. In language impairments, as well as in communicative exchanges occurring in multilingual contexts, morphological categories are not perfectly acquired. Empirical evidence relative to atypical contexts shows a peculiar asymmetry in the structure of language: alongside early development and strength of lexicon and phonology, late appearance and vulnerability of morphology. However, the weakness of morphology does not imply a minor role of this component into the grammar.

Keywords: faculty of language, morphology, language pathology, specific language impairment, language acquisition

1. Introduction

The vulnerability of the morphological component into the language architecture is the topic of this article. The primary evidence for such a vulnerability comes from language pathologies and language acquisition, two domains which entered into the agenda of linguists only recently, thanks to the cognitive drift impressed by the generative theory to the research on language.

The bulk of experimental work on language pathology has significantly expanded in the least decades. In particular, the so-called *Specific Language Impairment* (SLI) has been studied with reference to different natural languages (Bishop 2004). Other diseases related to language – such as for instance dyslexia, autism and deafness – have also been consistently investigated by linguists and psychologists.

The data collected in these different domains enlighten a common result, i.e. the vulnerability of morphology. Indeed, this component of language is often subject

to reduction and simplification processes, both in production and perception. Morphemes expressing categories such as number, gender, agreement, inflection and so on are often omitted or wrongly used by subjects affected by impairments concerning the faculty of language. The incomplete grammatical competence shown by subjects with atypical development seems to depend on a bad functioning of the neural circuits of memory, in particular of those implementing procedural memory.

On the other hand, in the process of first language acquisition, children start to speak by producing single words, i.e. signs mixing sound and meaning, and words become the milestones in constructing language, whereas morphology and syntax are acquired later (Guasti 2004). The same happens in second language acquisition in spontaneous contexts (Klein & Purdue 1997): phonology (often imperfect) and meanings are acquired first, while inflectional morphology is usually not learned by adults proficiently.

Once the special vulnerability of morphology has been recognised, the theoretical question on the background that linguists should face concerns the special place morphology occupies in the human faculty of language. In other terms, why is morphological competence selectively damaged in the atypical developments of language in the child? In parallel, in the light of the similar path shown in L1 acquisition as well as in L2 learning, what is the reason for the late competence in grammar with respect to phonology and lexicon? Are they special neural paths involved in the processing of morphology? Is morphology really necessary for language competence or it is only a secondary and subsidiary component? The first paragraphs of this paper will be devoted to reviewing the relevant literature relative to some language pathologies (§§ 2 and 3), whereas in the second part of the paper (§ 4) we will consider data coming from the research on Second Language acquisition as well as on multilingual contexts. A few considerations on language evolution and relative late appearance of grammar in human communicative code will close the article (§§ 5 and 6).

2.   Language pathologies

Among the pathologies concerning language, the *Specific Language Impairment* (SLI) has been one of the most studied diseases. As is well-known, the term *Specific Language Impairment* (henceforth, SLI) has been introduced by Leonard (1981). The selection of the word *specific* was not accidental, since it implicitly referred to the modularity of mind proposed by the Chomskian theory (cf. Chomsky 1988).[1]

---

1. On this challenged topic, see Fodor (1983, 2000), Marslen-Wilson & Tyler (1987), Carruthers (2006).

Indeed, SLI selectively affects the module responsible for language, without damaging other cognitive modules. Children affected by DSL show a reduced and lacking grammatical competence, both in terms of production and perception. SLI often shows comorbidity with other learning diseases, such as dyslexia and autism (Gooch et al. 2014). Moreover, it seems that SLI results from the combined influence of multiple genetic variants (Lai et al. 2001; Bishop 2002). In particular, the gene *Fox P2* (Gopnik & Crago 1991) has been considered for a while as the 'gene of grammar', although nowadays its role in the faculty of language has been downsized on the genetic, neurophysiological and psychological sides.

Considerable grammatical gaps are reported for children affected by SLI: they are not able to acquire the coding and decoding of inflection, agreement and other functional heads (Leonard 1998; Bishop 2004). However, not all the morphemes show the same behaviour; for instance, articles are normally used, whereas homophone clitic pronouns are not (Jakubowicz et al. 1998; Jakubowicz, 2003 for French; Bortolini et al. 1998; Bottari et al. 1998 for Italian).

Developmental Dyslexia (henceforth, DD) specifically refers to a disorder in learning to read in the context of normal intelligence.[2] Children affected by DD quite often show other impairments, such as dyscalculia, dyspraxia, SLI. Their phonological and then orthographic competence is insufficient, because not only can they not read fluently, but they are also unable to perform tasks requiring phonological awareness, such as identifying phonemes, building rhymes, or reproducing prosodic patterns (Bradley & Bryant 1983; Schankweiler et al. 1995; Snowling 2000). In other words, they are not able to play with the sounds of their mother tongue (Bryant 1995; Elbro 1996; Vellutino et al. 2004).

However, denomination tasks as well as morphosyntactic abilities are often impaired in dyslexic subjects (Nicolson & Fawcett 2008). Both denomination and reading are based on the recognition of forms and use the same neuronal circuits of sight (Dehaene & Cohen 2011). Both denomination and reading are acquired by the children in the first years of their life, and become early automatized cognitive behaviours, due to their frequent use. Moreover, these capacities, being automatic, are governed by the system of procedural (or implicit) memory (cf. the next paragraph). Therefore, we expect parallel difficulties in denomination and reading tasks in dyslexic children.

The test *Rapid Automatized Naming* (RAN; Denckla & Rudel 1976) is nowadays the best tool to discriminate between DD and SLI. As a matter of fact, dyslexic subjects achieve lower scores than the children affected by SLI as well as

---

2. *National Institute of Neurological Disorders and Stroke*, within the U.S. National Institutes of Health, retrieved 27 July 2016. Instead, in the case of *acquired dyslexia*, the loss of the ability to read is caused by brain damage.

by controls, because in the denomination tasks the lexicon is the only element involved, without any simultaneous activation of grammar (Bishop et al. 2009).

However, DD and SLI do not seem to be completely independent from one another. Instead, the two pathologies appear to share some elements. In particular, the empirical data collected in the last decades have shown relevant grammatical gaps in the linguistic competence of dyslexic subjects quite frequently.[3] The most involved areas concern morphological inflection, agreement, clitic pronouns, derivation and syntactic government. It is worth remarking that these grammatical categories are lately acquired by children with a typical linguistic development too, probably due to their major internal complexity.[4] Therefore, once more the analogy of behaviour between typical and atypical developments does not seem to be senseless. Moreover, the system of the procedural memory does not seem to work perfectly in children with SLI, once more suggesting a strong connection between the difficulties in managing grammar and reading.

It is also worth recalling that since the studies by Scarborough (1990, 1991) the evaluation of the morphological competence can be used as a diagnostic tool for predicting dyslexia before the beginning of elementary school. Experimental research in this domain has shown a distinct correlation between the development of grammar at preschool age and learning of reading at school.

Although scholars still discuss about the weight of the grammatical competence in the diagnosis of DD,[5] the relation between DD and SLI is nowadays one of the new frontiers of the research on atypical language development. A promising hypothesis might be to assume that both DD and SLI depend on more general and underlying gaps of cognitive nature, based on an anomalous or reduced activation of specific neuronal circuits crucially connected with the system of memory.

3.   Language pathologies and memory

The linguistic competence is strictly connected with the system of memory, in both its basilar components of *Long Term Memory* (LTM) and *Short Term Memory* (STM). In particular, the first component is traditionally divided into

---

3. For instance, Byrne 1981; Scarborough 1990, 1991; Gallagher et al. 2000; Waltzman & Cairns 2000 for English; Rispens 2004; Rispens & Been 2007 for Dutch; Fiorin 2010; Zachou et al. 2013; Guasti 2013; Guasti et al. 2015; Cardinaletti & Volpato 2015; Pivi & Del Puppo 2015; Marotta 2017, for Italian.

4. Here by complexity we mean degree of formality and abstractness.

5. See for instance Tallal & Piercy (1973); Bishop et al. (1999): Bishop (2004); Bishop & Snowling (2004).

two sub-systems, i.e. declarative (or explicit) memory and procedural (or implicit) memory. In the lexical selection, the declarative memory is activated, whereas the procedural memory is at work in the use of the grammatical rules.

According to the theoretical model of memory proposed by Ullman (2001, 2004), which is nowadays one of the most applied in psycholinguistic and behavioural studies, these two components of the LTM are in strict connection with the structure of grammar; in particular, the declarative memory is activated in the framework of lexical selection, whereas the procedural memory is activated in the use of the grammatical rules.

Subjects affected by SLI as well as by DD with agrammatism show a deficit of the procedural memory. Therefore, they are unable to properly control the combinatorial aspects of language, at all the levels, i.e. phonology, morphology and syntax. On the other hand, they appear to manage the lexicon, which is the component governed by the declarative memory. Recent neurophysiological studies using *neuroimaging* techniques have highlighted the relevance of the circuits located in Broca's area for the processing of language, not only at the level of motor control and in syntactic planning, but also in the activation of procedural memory.[6]

If DD is a language impairment involving not only phonology, but also grammar, i.e. morphology and syntax, how is it possible to discriminate between the two diseases? The relation of DD with SLI is indeed a widely debated topic in recent literature. According to some scholars (Tallal & Piercy 1973), DD would be a weak form of SLI. Evidence consists in the shared involvement of the system of procedural memory (as we have seen before) and the unidirectional connection between the two impairments: DD without SLI is found, not vice versa. However, neurobiological studies do non support their identity (Bishop & Snowling 2004).

A more relevant question appears to be that connected with the relation between competence and cognition. In other terms, are language pathologies completely dependent from the faculty of language or are they rather simple epiphenomena of a more general cognitive deficit?

4. Morphology in language acquisition

Longitudinal studies of the last decades devoted to language acquisition indicate a clear and constant path starting with early access to lexicon and phonology and later grammatical competence (Guasti 2007), regardless of the input received and the typology of L1 (with more or less inflection). In the case of spontaneous learning of a second language at adult age the steps are the same. The speakers of the

---

6. See the studies by Grodzinsky (2000, 2005); Grodzinsky & Amunts (2006).

so-called *Basic Variety* (Klein & Perdue, 1997) achieve a good lexical competence fairly soon alongside a more or less imperfect articulatory production; in parallel, their morphological and syntactic competence remains low. Derivation and inflection are poor, functional heads are often missing and the structure of the sentence is normally elementary. In the guided acquisition of L2 as well, both bound and free morphemes are often omitted, even in the most advanced steps. As far as Italian is concerned, the large amount of empirical data collected on the learning of Italian as L2 shows a constant instability of morphology, in the context of both spontaneous and guided learning (Giacalone Ramat 1986, 1988, 2003; Marotta & Meini 2012). The areas of greater difficulty are the inflection of nouns and verbs, agreement, prepositions and conjunctions, relative and passive clauses.

In a short paper published thirty years ago, Raffaele Simone (1988) focuses on morphology, underlining its vulnerability in a series of communicative contexts, such as pidgins and creoles, L2 spoken by the immigrants, child language, and deafness. In all these atypical contexts, the use of language shows a constant inhibition of the morphological features, suggesting a special degree of depth in the linguistic competence. Similarly, Anna Giacalone Ramat (2006) has underlined the special status of the morphological categories in the spontaneous acquisition of a second language at adult age. The reasons for such specificity would primarily be the arbitrary nature of morphological features as well as the typological distance between L1 and L2. Therefore, it is no surprise that a Chinese learner of Italian as L2 might have more problems in managing the verbal inflection of Italian than a French or German learner, since Chinese grammar is much more distant from Italian than in the case of German and French.

The special vulnerability of morphology is also observed in the languages spoken in multilingual contexts. The lack of inflectional morphology is indeed a typical feature of pidgins and creoles, which tend to be isolating languages. Although the latter may be considered the evolution of the former, on the pressure of Universal Grammar, according to Bickerton (1990), both *pidgins* and creoles are linguistic structures with a significantly reduced morphology and a quite elementary syntax; for instance, agreement is a very rare phenomenon.[7] The few morphemes surviving are normally derived by the source language and sometimes deprived of productivity, whereas the new ones are often governed by general principles of iconicity; e.g. reduplication of the noun to indicate plurality; addition of a temporal adverb to mark the past in the verbs, and so on. However, we are not claiming that pidgins and creoles are completely deprived of grammar: as shown by the data collected in the *Atlas of Pidgin and Creole Language Structures*

---

7. Cf. Seuren & Wekker (1986: 61): "It is generally agreed that creole languages have little or no morphology".

Online, relative to 76 languages around the world, some morphological features do occur in these languages.[8]

Coherent with the picture drawn so far, a strong trend towards the simplification of the potential morphology is observed in many substandard varieties occurring in many languages. A typical example is *African American Vernacular English* (AAVE), otherwise referred to as *Black American English* or *Ebonics*: the scarce morphology of English is drastically simplified in this case, the omission of the functional elements is normal (e.g. agreement between subject and verb, auxiliaries, clitics), whereas a trend to agglutination is reported (e.g. the use of <BIN>, standard English *been*, to express the past).[9]

After all, Meillet (1919) and Weinreich (1953) have already underlined the loss of inflectional marks in the case of linguistic contact, that is not only in pidgins and creoles but also in the context of diglossia as well as in bilingualism between L1 and L2.

Yet, what has to be remarked is the fact that all the processes of simplification of the grammatical structure we may observe in different contexts do not prevent or even block the communication among the speakers.

5.   Language evolution

Due to the dominant epistemological paradigm of Neo-Darwinism,[10] the topic of language evolution has once again entered the agenda of linguistics, after a long time of banning. The question is wrapped around this central nucleus: how did humans evolve so as to be able to learn language?

Scholars are aware of the arbitrariness intrinsically inherent to the topic, since there is no direct evidence for when and how our ancestors became able to speak. Nor is there evidence for what they could say: there are no fossil vowels or fossil words. Even the comparison to other species is not very helpful in this domain. Apes and monkeys do not have communicative codes similar to human language (Hedwig et al. 2015; Scott-Phillips 2016; Zuberbühler 2015). Nor do other species show a specific ability to acquire language in the sense of human grammar, even those with good equipment for vocal imitation (Fitch 2010).

---

8. See Michaelis (2013) and the website APICS online, <http://apics-online.info> (6 February 2018).

9. See the classic studies by Labov (1972) and Mufwene (2001).

10. The topic is marginal here; however, see Pinker & Bloom (1990); Hauser et al. (2002); Fitch et al. (2005); Pennisi & Falzone (2010); Banfi (2013).

In order to investigate language evolution in a plausible way, Jackendoff and Wittenberg (2017) have recently proposed 'reverse engineering'. The question that should be asked becomes then the following: if we admit that language traits did not all appear together, which components could have been useful in the absence of others?

The two authors draw a comparison with the evolution of human sight. If we consider its basilar components, i.e. retina and ocular muscles, we may ask if the retina could have been useful for vision even in the absence of muscles focusing the lens and moving the eyeballs or the muscles without a retina. We must assume that a primitive retina must have evolved before the lens and the muscles.

Language could have evolved in a similar way. The first step might be formed by a primitive system for communicating meaning via sound or gestures, without morphology or syntax. These latter components can improve a communication system, but they do not allow communication on their own. Only in a second step grammar appeared in the communication of human beings, in strict connection with the development of other cognitive capacities. With their words, "it makes sense that the connection between phonetics and meaning came first", followed by morphology and syntax as further refinements (Jackendoff & Wittenberg 2017: 220).

The same path is observed in the acquisition of the mother tongue: equipped with the language faculty, children start to communicate using words and sounds fairly soon, i.e. semantics and phonology, and only later do they begin to build up their grammar. This similarity appears to support a strong relation between ontogenesis and philogenesis, although in the case of children the path can be empirically demonstrated, whereas in the case of language evolution it can be speculated only. We do not have direct evidence on the birth of language in the human species before the coming of writing in human societies. The only evidence we have is indirect, relative to the anatomy of the skull alongside the reconstruction of some cultural aspects such as tools and draws.

Bickerton (1990, 2007) proposed two chronological steps in language evolution, i.e. *proto-language* and *modern language*, crucially distinct in relation to syntax processing. In the case of *proto-language*, syntax was simply reduced to word sequence linearly ordered according to general semantic principles,[11] whereas in *modern language* more complex hierarchies are developed. Adopting Bickerton's idea of *proto-language*, Jackendoff claims that some specific features already belonging to the *proto-language* still persist in the modern historical languages

---

11. A typical instance of such a proto-syntax is the principle *Agent First*, which is also usually employed in *pidgins* as well as in L1 and L2 acquisition (Jackendoff 2002; Jackendoff & Wittenberg 2017).

we know. These traces of the most ancient stages of language have the status of 'linguistic fossils' and may be found "in degraded forms of modern language" (Jackendoff 2002: 236).

Which are the degraded forms of language maintaining a sort of primitive form of communicative code deprived of grammar? First of all, just the disturbed contexts we have considered so far, that is language pathologies such as SLI or dyslexia, and agrammatic dementia. At the same time, evidence for preliminary states of human language could be the basic varieties in L2 adult learning, as well as pidgins and creoles. Second, onomatopoeias, interjections and other discourse signals, still occurring in natural languages, share a more transparent relation between sound and meaning, and are then closely related to the first step in language development (Jackendoff 2002: 240). The same lack of inflection occurring in some languages reaches the status of a linguistic fossil, since it could represent a sort of archaism in the evolution of language.

6. Concluding remarks

The critical review of data relative to different domains (clinical, acquisitional, sociolinguistic) and to different languages has allowed us to enlighten the special vulnerability of morphology. In language impairments, as well as in communicative exchanges occurring in multilingual contexts, morphological categories are not perfectly acquired.

In pathology, the faculty of language does not reach its complete maturation due to genetic and neurophysiological reasons. In the speech of adults learning a second language without an explicit teaching aimed at promoting metalinguistic awareness, the innate mechanisms for language acquisition are scarcely activated.

However, a limited grammatical competence does not inhibit communication: children affected by SLI or dyslexia may communicate with parents and other people in a satisfactory way. The same applies to speakers of L2 or a pidgin variety. Therefore we might challenge the relevance of morphology in language communication: human communication is saved even if morphology is reduced.

Indeed, human languages do not serve communication only. This is the main difference with respect to the communication codes used by the other animal species. Human languages also represent fine symbolic structures physiologically and cognitively constrained. And within such a structure morphology is not a kind of corollary, it is one of its basic components, one of the most defined traits of our species-specific code.

At the same time, we have to recognise that the morphological component can be more or less heavy within the general framework of a language. It is relatively

light in *pidgins* as well as in the spontaneous learning of L2 and in substandard varieties. It is also light in sign languages. It is variably light and incoherent in the speech of subjects affected by language pathologies. Still, morphology is somehow present in all natural languages, since it is one of their basic components. Take a language like Chinese as an example: morphological traits are very scarce, but they are not completely lacking.

Empirical evidence relative to different atypical contexts (crucially, language pathologies, spontaneous leaning of a second language) shows some asymmetry in the complex architecture of human languages: alongside early development and strength of lexicon and phonology, late appearance and vulnerability of morphology. Such an asymmetry does not imply a minor role of the morphological component into the structure of language, because there is no natural language that does not show at least some morphological features.

In conclusion, the vulnerability of morphology does not imply that it is useless or redundant. On the contrary, morphology still has the status of a fundamental nucleus of grammar, that is of a human symbolic structure.

Acknowledgment

References

Banfi, Emanuele (ed.). 2013. *Sull'origine del linguaggio e delle lingue storico-naturali*. Roma: Bulzoni.

Bickerton, Derek. 1990. *Language and Species*, Chicago: University of Chicago Press.

Bickerton, Derek. 2007. Language evolution: A brief guide for linguists. *Lingua* 117: 510–526. https://doi.org/10.1016/j.lingua.2005.02.006

Bishop, Dorothy V. M. 2002. The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders* 35(4): 311–328. https://doi.org/10.1016/S0021-9924(02)00087-4

Bishop, Dorothy V. M. 2004. Specific language impairment: Diagnostic dilemma. In *Classification of Developmental Language Disorders: Theoretical Issues and Clinical Implication*, Ludo Verhoeven & Hans van Balkom (eds), 309–326. Hillsdale NJ: Lawrence Erlbaum Associates.

Bishop, Dorothy V. M., Carlyon, Robert P., Deeks John M. & Bishop, Shawn J. 1999. Auditory temporal processing impairment: Neither necessary nor sufficient for causing language im-

pairment in children, *Journal of Speech, Language, and Hearing Research* 42: 1295–1310. https://doi.org/10.1044/jslhr.4206.1295

Bishop, Dorothy V. M. & Snowling, Margaret J. 2004. Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin* 130(6): 858–886. https://doi.org/10.1037/0033-2909.130.6.858

Bortolini, Umberta, Arfé, Barbara, Caselli, Maria Cristina, Degasperi, Luisa, Deevy, Patricia & Leonard, Laurence B. 2006. Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language and Communication Disorders* 41: 695–712. https://doi.org/10.1080/13682820600570831

Bortolini, Umberta, Leonard, Laurence B. & Caselli, Maria Cristina. 1998. Specific language impairment in Italian and English: Evaluating alternative accounts of grammatical deficits, *Language and Cognitive Processes* 13(1): 1–20. https://doi.org/10.1080/016909698386573

Bottari, Piero, Cipriani, Paola, Chilosi, Anna Maria & Pfanner, Lucia. 1998. The determiner system in a group of Italian children with SLI. *Language Acquisition* 7(2): 285–315. https://doi.org/10.1207/s15327817la0702-4_7

Bradley, L. & P. Bryant. 1983. Categorizing sounds and learning to read. A causal connection. *Nature* 301: 419–421.

Bryant, P. 1995. Phonological and grammatical skills in learning to read. In *Speech and Reading. A Comparative Approach*, Beatrice de Gelder & José Morais(eds). Hillsdale NJ: Lawrence Erlbaum Associates.

Byrne, Brian. 1981. Deficient syntactic control in poor reader. Is a week phonetic memory code responsible? *Applied Psycholinguistics* 2: 201–212. https://doi.org/10.1017/S0142716400006512

Cardinaletti, Anna & Volpato, Francesca. 2015. On the comprehension and production of passive sentences and relative clauses by Italian university students with dyslexia. In *Structure, Strategies, and Beyond: Studies in Honour of Adriana Belletti* [Linguistik Aktuell/Linguistics Today 223, Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds), 279–301. Amsterdam: John Benjamins. https://doi.org/10.1075/la.223.13car

Carruthers, Peter. 2006. *The Architecture of the Mind*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199207077.001.0001

Chomsky, Noam. 1988, *Language and Problems of Knowledge*. Cambridge MA: The MIT Press.

Dehaene, Stanislas & Cohen, Laurent. 2011. The unique role of the visual word form area in reading. *Trends in Cognitive Sciences* 15(6): 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Denckla, Martha Bridge & Rudel, Rita G. 1976. Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia* 14(4): 471–479. https://doi.org/10.1016/0028-3932(76)90075-0

Elbro, Carsten. 1996. Early linguistic abilities and reading development: A review and a hypothesis about underlying differences in distinctiveness of phonological representations of lexical items. *Reading and Writing* 8: 453–485. https://doi.org/10.1007/BF00577023

Fiorin, Gaetano. 2010. *Meaning and Dyslexia: A Study on Pronouns, Aspect, and Quantification*. Utrecht: LOT.

Fitch, W. Tecumseh. 2010. *The Evolution of Language*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511817779

Fitch, W. Tecumseh, Hauser, Marc D. & Chomsky, Noam. 2005. The evolution of language faculty: Clarifications and implications. *Cognition* 97: 179–210. https://doi.org/10.1016/j.cognition.2005.02.005

Fodor, Jerry A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge MA: The MIT Press.

Fodor, Jerry A. 2000. *The Mind Doesn't Work That Way*. Cambridge MA: DThe MIT Press.

Gallagher, Alison M., Frith, Uta & Snowling, Margaret J. 2000. Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatrics* 41: 203–211. https://doi.org/10.1017/S0021963099005284

Giacalone Ramat, Anna (ed.). 1986. *L'apprendimento spontaneo di una seconda lingua*. Bologna: il Mulino.

Giacalone Ramat, Anna (ed.). 1988. *L'italiano tra le altre lingue. Strategie di acquisizione*. Bologna: il Mulino.

Giacalone Ramat, Anna (ed.). 2003. *Verso l'italiano. Strategie di acquisizione*. Roma: Carocci.

Giacalone Ramat, Anna. 2006. Acquisizione della morfologia di italiano L2. In *Zhi. Scritti in onore di Emanuele Banfi*, Nicola Grandi & Gabriele Iannaccaro (eds). Roma: Caissa.

Gooch, Debbie, Hulme, Charles, Nash, Hannah M. & Snowling, Margaret J. 2014. Comorbidities in preschool children at family risk of dyslexia. *Journal of Child Psychology and Psychiatry* 55/3, 237–246. https://doi.org/10.1111/jcpp.12139

Gopnik Myrna & Crago, Martha B. 1991, Familial aggregation of a developmental language disorder. *Cognition* 39: 1–50. https://doi.org/10.1016/0010-0277(91)90058-C

Grodzinsky, Yosef. 2000. The neurology of syntax. *Behavioral and Brain Sciences* 23: 1–71. https://doi.org/10.1017/S0140525X00002399

Grodzinsky, Yosef. 2005. Syntactic dependencies as memorized sequences in the brain. *Canadian Journal of Linguistics* 50: 241–266. https://doi.org/10.1017/S0008413100003728

Grodzinsky, Yosef & Amunts, Katrin. 2006. *Broca's Region*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780195177640.001.0001

Guasti, Maria Teresa. 2004. *Language Acquisition. The Growth of Grammar*. Cambridge MA: The MIT Press.

Guasti, Maria Teresa. 2013. Oral skills deficit in children with Developmental Dyslexia. In*Advances in Language Acquisition*, Stavroula Stavrakaki, Marina Lalioti & Polyxini Konstantinopoulou (eds), 416–424. Newcastle upon Tyne: Cambridge Scholars.

Guasti, Maria Teresa, Branchini, Chiara, Vernice, Mirta, Barbieri, Lina & Arosio, Fabrizio. 2015. Language disorders in children with Developmental Dyslexia. In *Specific Language Impairment. Current Trends in Research* [Language Acquisition and Language Disorders 58], Stavroula Stavrakaki (ed.), 35–55. Amsterdam: John Benjamins. https://doi.org/10.1075/lald.58.02gua

Hauser, Marc, Chomsky, Noam & Fitch, W. Tecumseh. 2002. The faculty of language: What is it, Who has it, and How did it Evolve? *Science* 298: 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Hedwig, Daniela, Mundry, Roger, Robbins, Martha M. & Boesch, Christoph. 2015. Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition* 18: 423–435. https://doi.org/10.1007/s10071-014-0812-6

Jackendoff, Ray. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3(7): 272–279. https://doi.org/10.1016/S1364-6613(99)01333-9

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198270126.001.0001

Jackendoff, Ray. 2013. Constructions in the parallel architecture. In *The Oxford Handbook of Construction Grammar*, Thomas Hoffmann & Graeme Trousdale (eds), 70–92. Oxford: OUP.

Jackendoff, Ray & Vitteberg, Eva. 2017. Linear grammar as a possible stepping-stone in the evolution of language. *Psychonomic Bulletin and Review* 24: 219–224.
https://doi.org/10.3758/s13423-016-1073-y

Jakubowicz, Celia. 2003. Computational complexity and the acquisition of functional categories by French-speaking children with SLI. *Linguistics* 41(2): 175–211.
https://doi.org/10.1515/ling.2003.007

Jakubowicz, Celia, Nash, Léa, Rigaut, Catherine & Gérard, Cristohphe-Loic. 1998. Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition* 7: 113–160.
https://doi.org/10.1207/s15327817la0702-4_3

Klein, Wolfgang & Perdue, Clive. 1997. The basic variety (or: Couldn't natural languages be much simpler? *Second Language Research* 13(4): 301–347.
https://doi.org/10.1191/026765897666879396

Labov, William. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia PA: University of Pennsylvania Press.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413: 519–523.
https://doi.org/10.1038/35097076

Leonard, Laurence B. 1981. Facilitating linguistic skills in children with Specific Language Impairment, *Applied Psycholinguistics* 2(2): 89–118.
https://doi.org/10.1017/S0142716400000886

Leonard, Laurence B. 1998. *Children with Specific Language Impairment*. Cambridge MA: The MIT Press.

Marotta, Giovanna. 2017. Dislessia, grammatica e dintorni. In *Lingua e patologia. Le frontiere interdisciplinari del linguaggio*, Francesca M. Dovetto (ed.), 85–113. Roma: Aracne.

Marotta, Giovanna & Meini. 2012. Spatial Prepositions in Italian L2. In *Space and Time in Languages and Cultures: Linguistic Diversity* [Human Cognitive Processing 36], Luna Filipovic & Katarzyna M. Jaszczolt (eds), 289–324. Amsterdam: John Benjamins.
https://doi.org/10.1075/hcp.36.16mar

Marslen-Wilson, William & Tyler, Lorraine Komisarjevsky. 1987. Against modularity. In *Modularity in Knowledge Representation and Natural Language Understanding*, Jay L. Garfield (ed.). Cambridge MA: The MIT Press.

Meillet, Antoine. 1919. Le genre grammatical et l'élimination de la flexion. *Scientia* 25. Reprinted in Idem, *Linguistique historique et linguistique générale*, Vol. I, 1–18. Paris: Champion.

Michaelis, Susanne M., Maurer, Philippe, Haspelmath, Martin & Huber, Magnus (eds). 2013. *The Atlas of Pidgin and Creole Language Structures*. Oxford: OUP.

Mufwene, Salikoko S. 2001. What is African American English?, In *Sociocultural and Historical Contexts of African American English* [Varieties of English Around the World G27], Sonja L. Lanehart (ed.), 21–52. Amsterdam: John Benjamins.
https://doi.org/10.1075/veaw.g27.06muf

Nicolson, Roderick I. & Fawcett, Angela J. 2008. *Dyslexia, Learning, and the Brain*. Cambridge MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262140997.001.0001

Pennisi, Antonio & Falzone, Allesandra. 2010. *Il prezzo del linguaggio. Evoluzione ed estinzione nelle scienze cognitive*. Bologna: il Mulino.

Pinker, Steven & Bloom, Paul. 1990. Natural language and natural selection, *Behavioral and Brain Sciences* 13: 707–784. https://doi.org/10.1017/S0140525X00081061

Pivi, Margherita & Del Puppo, Giorgia. 2015. L'acquisizione delle frasi relative restrittive in bambini italiani con sviluppo tipico e con dislessia evolutiva. In *Grammatica applicata: Apprendimento, patologie, insegnamento*, Maria Elena Favilla & Elena Nuzzo (eds), 59–73. Milano: Officinaventuno.

Rispens, Judith E. 2004. Syntactic and Phonological Processing in Developmental Dyslexia. PhD dissertation, University of Groningen.

Rispens, Judith E. & Been, Pieter. 2007. Subject-verb agreement and phonological processing in developmental dyslexia and SLI: A closer look. *International Journal of Language and Communication Disorders* 42(3): 293–305. https://doi.org/10.1080/13682820600988777

Scarborough, Hollis S. 1990. Very early language deficits in dyslexic children. *Child Development* 61: 1728–1743. https://doi.org/10.2307/1130834

Scarborough, Hollis S. 1991. Antecedents to reading disability: Preschool language development and literacy experiences of children from dyslexic families. *Reading and Writing: An Interdisciplinary Journal* 3: 219–233. https://doi.org/10.1007/BF00354959

Seuren, Pieter & Wekker, Herman. 1986. Semantic transparency as a factor in creole genesis. In *Substrata Versus Universals in Creole Genesis*, Muysken and Smith (eds), 57–70, Amsterdam: John Benjamins. https://doi.org/10.1075/cll.1.05seu

Scott-Phillips, Thomas C. 2016. Meaning in great ape communication: Summarising the debate. *Animal Cognition* 19: 233–238. https://doi.org/10.1007/s10071-015-0936-3

Simone, R. 1988. "Fragilità" della morfologia e "contesti turbati". In *L'italiano tra le altre lingue. Strategie di acquisizione*, A. Giacalone Ramat (cur.), 91–98, Bologna: il Mulino.

Snowling, Margaret J. 2000. *Dyslexia*. Oxford: Blackwell.

Tallal, P. & Piercy, M. 1973. Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241: 468–469. https://doi.org/10.1038/241468a0

Ullman, Michael T. 2001. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research* 30(1): 37–69. https://doi.org/10.1023/A:1005204207369

Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92: 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Vellutino, Frank R., Fletcher, Jack M., Snowling, Margaret J. & Scanlon, Donna M. 2004. Specific Reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry* 45: 2–40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Waltzman, Dava E. & Cairns, Helen S. 2000. Grammatical knowledge of third grade good and poor readers. *Applied Psycholinguistics* 21: 263–284.
https://doi.org/10.1017/S014271640000206X

Weinreich, Uriel. 1953. *Language in Contact. Findings and Problems*. New York NY: Linguistic Circle of New York.

Zachou, Angeliki, Partesana, Enrica, Tenca, Emanuela & Guasti, Maria Teresa. 2013. Production and comprehension of direct object clitics and definite articles by Italian children with Developmental Dyslexia. In *Advances in Language Acquisition*, Stavroula Stavrakaki, Marina Lalioti & Polyxeni Kostantinopoulou (eds), 464–471, Newcastle upon Tyne: Cambridge Scholars.

Zuberbühler, Klaus. 2015. Linguistic capacity of non-human animals. *Wiley Interdisciplinary Reviews: Cognitive Science* 6: 313–321.

CHAPTER 26

Some Celto-Albanian isoglosses
and their implications

John Trumper
University of Calabria

Çabej (1969) originally posited a small number of Albanian-Celtic-Germanic
isoglosses. Apart from a more detailed discussion of *besa* and *njerí*, we highlight
in greater depth the areal diffusion of the I-E diffusion of Albanoid *bërrakë,
e blertë, brī* and *dritë*. We add in-depth observations on the Celto-Albanian
binomial *ardracht* (Old Irish) – *dritë* (Albanoid), where the Celtic terms involved
are traceable to Gaulish *dercos* and *uodercos* of texts and inscriptions. Such
isoglosses are developed here and tend to substantiate Hamp's view that the
Albanoid Urheimat was originally Central-North Europe rather than its present
day Mediterranean Heimat.

Keywords: Albanian, Celtic, Germanic, lexical isoglosses

1.   Celto-Albanian isoglosses

Çabej (1969) is perhaps the first systematic attempt to come to grips with the prob-
lem of possible Germano-Albanian and Celto-Albanian isoglosses. The solution of
these would help to establish an Urheimat for Proto-Albanoid groups in Central
Europe, north of their actual territory. Originally, Jokl (1923) had underpinned
isoglosses between Albanoid and Baltic languages, Greek and Armenian, but had
overlooked possible relations with Celtic and Germanic. The only case not dealt
with by Çabej in those discussed below is that of *Bri*, which he had hoped to re-
solve in a later moment (Çabej 1969: 186).[1] Huld (1984) criticized these first stud-
ies, rejecting all the lexical isoglosses except *drekë, dele*, and *shpen[d]*, overall a
rather facile criticism, for motives I shall try to elaborate. Trumper (1999, 2002)

---

1. "Hoffe ich an anderer Stelle zu handeln".

had suggested other cases, such as *besa*, a central topic in Albanian culture.[2] There is an obvious connection with Old Irish *bés[s]*, of similar meaning, < Old Irish or Late Gaulish *bessu* (Virgil Grammaticus). The problem here is that of a possible double origin, either from *BEND-TUH- 'binding' or from *BEID-TUH- 'persuasion', 'confidence' etc., so that Pokorny in IEW 117 and IEW 127 (1959) gave a double listing for Albanian *besë –a*, without treating Celtic cognates. Both Old Irish *béss* and Middle Breton *boas* (Ernault 1895: 72) are comparable outcomes.[3] In fact both Hamp's considerations and Çabej (1982: 110) rightly conclude that a reasoned choice between *BEND-TUH- and *BEID-TUH- is not possible, even though Meyer (1891: 33) had tried to argue in favour of the first, Huld (1984: 143) in favour of the second (< *BID-TIH- a reduction of *BEID-TIH-), without much logic or success. LEIA B–43 seemed to think that all British Celtic forms were incomparable ("les formes brittoniques sont en tout cas aberrantes"), though only Welsh *moes* seems to have the wrong vocalic outcome (it seems to suggest a Proto-Celtic *BAISS- not *BEISS-). Whichever of the two possible etyma is appropriate, it cannot be denied that there is an undoubted relationship between Old Irish *béss*, Middle Breton *boas* and Albanian *besë*. More doubtful is the relation of Welsh *moes* with all these, notwithstanding GPC's comments, probably based on Stokes-Bezzenberger (1894: 174)'s original hypothesis (base *bêssu- as in Virgilius Grammaticus > Irish *bés*, "cymrisch *moes* aus *boes*" etc.). Pisani (1959) had also tried to establish more lexical relations between Albanian and North Europe Indo-European, even if occasionally ignoring macroscopic cases like *njer, njerí*, forgetting Celtic outcomes of the IE base *H$_a$NER- (rather more than those usually considered). Apart from the cases in the poetical sections of the Llyfr Coch Hergest quoted for the early thirteenth century in GPC 2571A-B, we find the Middle Welsh couplet *ener/ ner*, in the same period, in the poetry of the bard Cynddelw 16. 206, as coupled *ener/ ner* in 26. 40–41,[4] in the mid twelve hundreds in the bardic poetry of Meilyr's sons Gwalchmai (9. 120) and Elidir Sais (17. 46), of Gwalchmai's son Einion (28. 40), and of the bard Dafydd Benfras (26. 49) "O wythlonder, ner, nerth a gedwy" (in fierceness a prince, [your] strength holds still strong), (27. 28)

---

2. It is present in the first Italo-Albanian texts, e.g. Matranga's *Dottrina Cristiana* 1592, folio 30 v. 8: see Sciambra (1964: 58).

3. Irish *bés* could well be from an earlier *beiss, Breton *boas* could be the outcome of either *beiss or baiss (Proto-Celtic *ei and *ɛi < *ai merge in Breton and Cornish but not in Welsh, which presents *ei > 'wy', *ai > *ɛi > 'oe').

4. 26. 40–41 (page 327) "Gwas a las o Leissyaun ener/ Llary Ennia6n, llyw llwyrdaun llwr6 ner" (A youthful prince who slew the Lleision, / generous Einion, a leader, perfect his gifts as prince).

"Och am ner, muner mynogrwydd" (lament for a prince, a noble lord),[5] in the mediaeval Welsh translation of the *Parvum officium beatæ Mariæ V.* (*Gwasanaeth Meir*, matins 10. 26[6]), up until Dafydd ap Gwilym in the thirteen hundreds (poem 5. 17), thence up to modern usage.[7] It would not appear to occur in Old Welsh, though the Old Welsh section of the *Canu Taliesin* (Williams's edition) VIII. 3 has "gweleis i rac **neb** nym gweles/ pop annwyl. ef diwyl y neges" (*neb* for *nep* would be strange in Old Welsh![8]), which would have more sense if the reading were "gweleis i rac **ner** nym gweles/ pop annwyl. ef diwyl y neges". One might also add, from the Old Welsh sections of the Aneirin, Williams' corrective reading of line 381 as *budic ener enhy* (victorious [was] a brave prince) instead of the incomprehensible reading *budic e ren eny* of Evans' diplomatic edition. Thus the Old Welsh use of *Ner/ Ener* is directly comparable with Old Irish *nár* 'generous, magnanimous' as princely attributes (with secondary lengthening), and the metaphorical use as *ner* 'wild boar' (Gamkrelidze-Ivanov 1995: 158, 703; Mallory & Adams 1997: 366). The comments in Loth (1924, 207–208) on *Ner* and its IE relations, adding Breton *nerein* "fortify" (Vannetais dialect), still hold good.[9] The *E-* in *Ener*, like Greek Ἀ- in ἀνήρ, supports the hypothesis of a word-initial laryngeal, the whole an areality that links Albanian *njerí* not only with Greek and Armenian (as well as Indo-Iranic) but also with Celtic congeners.

Central cases to the argument, apart from Albanoid outcomes of *BEID-TU–/ *BEND-TU- and H$_a$NER-, which I propose discussing, are the following.

---

5. See also 29. 24, 29. 100, 31. 22, 35. 4 etc.

6. This seems to be the first religious use of *Ner* as Lord = God (literally *Nef Ner* 'Lord of Heaven').

7. Up to the use of *Ener* 'Lord' in post-1600 Welsh Protestant hymnology.

8. Williams is following Evans (1910)'s diplomatic version, v. 18–19, page 62 "g6eleiſ i rac neb nym gweles/ pop ann6yl, ef diwyl y neges". He admits, in his note, however, that in this orthography 'p' and 'r' are difficult to distinguish and concludes "barnaf mai ner oedd y gwreiddiol yma, 'o flaen (yr) arglwydd'…" (I judge that *ner* was the original [reading] here: 'in front of/ before the prince'), comparing the sense with line 14 of the same poem where we find 'glyw', lord ("yscawydawr y rac glyw gloyw glas gwen": shield before a splendid lord with [his] grey smile).

9. Fleuriot (1964: 266) had already linked Welsh and Breton forms with Irish ner, near, while Schrijver (1995: 350) had already postulated the word-initial laryngeal of H$_2$NER-, H$_2$NER-TO- for the Celtic outcomes.

2.   *Bërrakë -a* 'bog'

Bërrakë -a 'bog' (Meyer 1891: 39), Italo-Albanian Birrakë (Giordano 1963: 31, 33; Raimundi 2001: 271; Baffa 2009: 23). Meyer (1891) proposed as origin the IE colour adjective base *BLEI-DRO-, from IEW 155–156. Huld (1984: 45–46) accepted this hypothesis, which I would exclude on semantic grounds. I would also exclude Orel (1998: 23)'s derivation *bërrakë < birë* 'cavity' both on formal (phonological) and semantic grounds. Çabej (1969: 176) and Çabej (1976: 215–217; 1981: 214, 254; 1987: 232) offered two hypotheses: (1) a Turkish loan (< *bare, biral*), (2) a Western IE isogloss, starting from Hesychius' βραγός· ἔλος, of Galatian origin. He considered similar Slav terms congeners. Skok 1. 109–110 thought the Albanian, as well as late Greek μπάρα, borrowings from Slav (Meyer's first hypothesis), which he proposed as co-radicals, without the theme's being doubled, of Sanskrit barburá and Greek βόρβορος < IEW 482 *GʷOR(-GʷOR-). Had Skok been on the right track, then one would have to suppose Southern Slav congeners to have been Greek borrowings (< βόρβορος). Were one to follow the Turkish hypothesis, the only solution would be in terms of Starostin- Dybo- Mudrak 2. 904–05 Proto-Altaic *MĀRO- > Proto-Turkish *BOR- 'marsh; muddy terrain' etc. However, an –AK- formant would seem decidedly more IE than Altaic. In IE terms *MRK- > *MERK-, *MREK- (Pokorny 739–740) (1) 'rot; putrefy', (2) 'putrid water', (3) 'marshland', is well represented in Celtic, both base and derivatives, e.g. *MRK-NO- > *mrakno- > *bragno- in Welsh (braenar etc.), Old Cornish (brein), Old Breton (arcibrenou), whence Middle Irish brén etc.[10] It has a constant presence throughout Celtic, from Gaulish to the modern languages. Consequently, there is a case for a Celto-Albanoid isogloss based on the *MRK- root, if we hypothesize that Albanian bërrakë can be taken to this base, even though the -AK- formant has all the appearance of a remote Celtic borrowing.

3.   *Blertë* 'greenish; verdant'

Blertë 'greenish; verdant'. The adjective is also present in Italo-Albanian (Giordano 1963: 35). Meyer (1891: 39) *bl'eroń* originally proposed starting from a *blerë* adjective comparable with Old Slav *blědъ* (Berneker 1908–1913: 60). Orel (1998: 29) held that the presence of *e blerë* in Italo-Albanian was sufficient proof of its primacy, defending Çabej's hypothesis of a Western European *cum* Albanian isogloss. Çabej (1976: 260–261), as Meyer (1891), took these outcomes to

---

10. The Celtic passed as *bragnicare* into Late Latin, with further Gallo-Romance outcomes. Cp. observations in Jud (1926).

*BLE[H]I-D-RO- > Proto-Albanoid *bleð-rë, congener of Balto-Slav terms. Huld (1984: 45) and Demiraj (1997: 104–105) followed the same line of argument. Çabej, later followed by Orel, noted the possibility of relating all these not only with Latin flāuus and flōrus,[11] but also with Germanic (Old Norse *blár*) and Celtic (Middle Irish *blár*, Middle/ Modern Welsh *blawr*, cp. Old Irish *blá*), in other words the possibility of setting up an important lexical isogloss between Albanoid and Western European IE languages. To go further back in time, we note that Thomas (1916) had already commented Old Provençal *blar* and Old French *bler*, *blair* (not only the derivative *blaireau*)[12] in the 12th–13th centuries (taken up by Von Wartburg in FEW 1. 401), while Delamarre (2001: 67) and Lambert (2003: 191) discussed the probable presence in Gaulish of a Blāros theme. Von Wartburg op. cit. was convinced that Germanic reflexes were also of Celtic origin.[13] Developments of the type IE *BLOH-RO- (underlying the Celtic forms) would normally give Proto-Albanoid *blōr- (+ – to-) > e blerë, e blertë, so there is no difficulty in setting up a Celto-Albanian lexical isogloss in this case.

4.   *Brī briri*, Gheg *brî brini* 'horn'

Brī briri, Gheg *brî brini* 'horn'. The word is also, whether as *brī briri* or *brī briu*, Ital-Albanian (Scutari 1991: 72 *brri –u*; Pignoli & Tartaglione 2007: 16 *brri*; Baffa 2009 *bri –u*). Çabej (1976: 321–23; 1982: 164) commented possible relationship with Classical Greek βρέντος, βρένδος (commented on in Hesychius), hence with Scandinavian (N. Germanic) and Baltic forms: he even suggests connections with a Gaulish *brīno- 'baguette' (see FEW 1. 528–531). Demiraj (1997: 110–111) proposed two hypotheses, without any decisive choice between them: (1) IEW 172 *BRUH-, (2) a dual *B[E]R-IH- 'Pfeil; Bogen'. Porzig (1974: 210) had considered this outcome merely Baltic, Germanic (essentially Scandinavian), and Illyrian, but we note that correspondences are not only Balto-Germanic-Albanoid but also Celtic. Compare Goidelic *braine*, *broine*, still better Old Welsh Ninth Century *breni* in "ír bréní. í. proram" (Juvencus 31, Stokes 1865: 399, comments in Falileyev 2000: 18), 10th–11th century Old Cornish *brenni* (Campanile 1974: 17, "brenniat .i. proreta"), Breni in Welsh place naming (1100, in the Breuddwyd Macsen Wledig, Llyfr Gwyn Rhydderch 94b1–4). Rather than FEW's *brīno- I propose

---

11. Demiraj (1997: 104 ff.) commented that the relation was problematic, especially if compared with Oscan *Fluúsai* and outcomes of the remote IE base *BLOH-.

12. See also Jud (1926) on the Gaulish elements in Romance.

13. Op. cit. "Neben dem kelt. Kommt auch das germ. in betracht für das etym. Ndl. blaar …".

starting from a Proto-Form *brĕno- (1) prominent part, (2) promontory, (3) forehead. IEW 167 *BREN–/ *BRON- collates Celtic outcomes with Latin frōns, frŏntem, even with Scandinavian forms.[14] It might well be the case to join up IEW 167 *BREN–/ *BRON- with IEW 168–69 *BREN-TO- (Greek, Albanoid, Italic, Germanic and Baltic), meaning 'curved' (> 'horn') and '(curved) prominence' > (1) 'prow', (2) 'promontory'. If this solution proves possible, then one can, without any difficulty, create a complex lexical isogloss Baltic-Germanic-Celtic-Italic-Albanoid.

5. *Dritë –a* 'light'

Dritë –a 'light', the most interesting case. The word is also Italo-Albanian (Giordano 1963: 83; Scutari 1991: 144; Pignoli & Tartaglione 2007: 36; Baffa 2009: 31). Meyer (1891: 749) had already hypothesized an outcome of IEW 213 *DERK'- > *DRK'-TI-, with *ditë* 'day' as a contaminating element ("*Verwechselung mit mir dit*"). Jokl, in his contributions, supposed only *DRGTI-, comparing with Indo-Iranian lexical choices, as did Çabej (1982: 213, 244, 247, 252) and Çabej (1987: 111, 252). Porzig (1974: 149) adds no further details. Later Orel (1998: 75) (*dritë*) showed indecision between IE *DERK'- and IE *GERH- 'shine; brighten', the remote origin of Old Church Slavonic zьrěti. Demiraj (1997: 145) objected that it was formally impossible to derive from this second IE base, turning back to a series of derivations IE *DERK'- > *DRK'-TI- (Albanian and Indo-Iranian), > *DRK'- and *DRK'-SI- (Celtic). Celtic forms are Old Irish *derc, dercaid, airdirc*, Early Middle Welsh *ardderchog, drych* (*drychaf*),[15] Old Breton *derc'h, erderc'h*, as well as in 200 AD Gaulish DERCOS 'eye' used in an apparently meaningless charm (in lieu of any possible medical cure) by Marcellus Empiricus (Antoninus Pius' personal doctor) in his *De medicina* VIII. 172.[16] Gaulish personal names such as Dercus

---

14. Note that Old Norse *brandr*, both 'ship's prow' and 'front of the house', might be considered a Celtic loan, a less possible hypothesis in the case of Old English *brant*, Old Norse *brattr*.

15. From a base *EK[S]-DRK'-SI- we have the common *edrych* 'glance', *edrychaf* 'look at', distinct from *drychaf* 'mirror' (verb). Welsh also presents derivatives from a morphologically more complex *DRK'-S-MA- in *drem, dremaf*. For the history of all such forms, see Delamarre (2001: 116) and Lambert (2003: 180).

16. "Ter carmen hoc dices et totiens spues: INMON DERCO MARCOS AXATISON". The charm is used 'ad oculorum causas remedia physica'. Fleuriot (1974) and Lambert (2003) divided the phrase more appropriately in IN MON DERCO [A]MARCOS AXAT ISON 'from [lit. in] my eye, Amarcus, take away this thing', where *Amarcus* is a divinity concerned with sight

and Dercoiedus, from the same base, also exist documented. Diffusion of the IE forms is, then, as follows:

*T'RK[']-: Indo-Iranian; Greek; Celtic; (perhaps Italic[17]); *T'RK'-S-: Indo-Iranian; Celtic; (doubtfully Germanic);

*T'RK'-S-M-: Greek; Celtic;

*T'RK'-T-: Albanoid [*dritë*]; Celtic: Irish *andracht/ indrocht* ,'dark' < 'non-light', or even *ardracht* 'bright; solar': with morphology *N-T'[E]RK'-T-, *[P]AR[E]-T'[E]RK'-T-.

Most of the areal diffusion shown occurs in peripheral IE areas (Indo-Iranian, Greek, Celtic), but it is important to note that the fourth type of development (*T'RK'-T-) is clearly documented in Albanian and Celtic. LEIA A−76/ A−87 demonstrates unequivocally that Irish *andracht* shows a morphological structure *N-T'RK-TO-, while *ardracht* a structure *[P]AR[E]-T'RK-TO-. The Albanoid-Celtic isogloss is thus perfect.

6.   Conclusions

We have attempted to show (1) that, as Çabej argued, there are definite lexical isoglosses that unite Albanian with Celtic, both with Germanic. These would suggest an Albanoid Urheimat in a more northern point of Central Europe than the present-day South European Heimat (E. P. Hamp's hypothesis, cf. Hamp 1966, 1972, 1996). (2) It is too simplistic to attribute too many Albanian phenomena to a straightforward and simple Roman influence during the Empire and to the importance of the Via Egnatia. Overall, the complex of Albanian dialects remains a solid block of the Albanoid group still relatable with Messapic (observed in place naming in Apulia: some towns have no etymon outside Albanoid sources, for example in toponyms such as Manduria[18]). Any label such as 'Illyrian' only creates embarrassment for its conceptual paucity. Albanoid is perhaps a better term, referring to a specific ethnolinguistically pertinent and historically compact language group,

---

(cp. Old Irish *amharc*) and *axat* a -SA- subjunctive of the basic verb root *H$_a$EG- (*agsad* with a similar morphology to Latin *faxat*).

17. That is, if we take the Old Umbrian (Italic) verb *terka-* to be an outcome of S-DERK- (Tavole Iguvine III. 9 "*Sakre uuem uhtur teitu puntes terkantur*" [To the Quinion gods, oh sacrificer, that they may see the sheep as a sacrificial offering]). Unless, of course, this is not just a banal graphical substitution t *pro* d.

18. Comprehensible only as a compounding of *mënd* (see Albanian *mëz/mâz*) and *urë* 'bridge', meaning 'bridge of the heifers', with an obvious reference to a central cattle mart. There are numerous other examples.

as Mallory (1999: 76) had stated, of "linguistically related tribal groups which we must somewhat uncomfortably label as Illyrian". We try not to be uncomfortable.

Caro Leonardo, semper ad astra!

Abbreviations

DIL	see Royal Irish Academy 1990
FEW	see von Wartburg (1929)
GPC	see Thomas (1950–2000)
IE	Indo-European
IEW	see Pokorny (1959)
LEIA	see Vendryes-Bachellery-Lambert

Acknowledgments

Literary Works

Berneker, Erich. (1908–1913). *Slavisches etymologisches Wörterbuch*. Heidelberg, Winter.

Costigan (Bosco), Nora G. (ed.). 1995. *Gwaith Dafydd Benfras ac Eraill o Feirdd Hanner Cyntaf y Drydedd Ganrif ar Ddeg*. Cardiff : University of Wales Press.

Evans, John G. (ed.). 1910. *The Text of the Book of Taliesin*. Llanbedrog: By the author/editor.

Evans, John G. (ed.). 1911. *The Poetry in the Red Book of Hergest*. Llanbedrog: By the author/editor.

Evans, John G. & Jones, Robert M. 1973. *Llyfr Gwyn Rhydderch. Y Chwedlau a'r Rhamantau.*. Gwasg Prifysgol Cymru: Caerdydd.

Jones, Nerys Ann & Parry Owen, Ann (eds). 1991. *Gwaith Cynddelw Brydydd Mawr*. Cardiff: University of Wales Press.

Jud, Jakob. 1926. *Mots d'origine gauloise*, 4e série. ROMANIA, pp. 328–348.

Parry, Thomas (ed.). 1979. *Gwaith Dafydd ap Gwilym*. Cardiff: University of Wales Press.

Raimundi, Filomena. 2001. *Rivista dell'Istituto Lombardo di Arti e Scienze XLI*: 211–270.

Roberts, Brynley F. (ed.). 1961. *Gwasanaeth Meir*. Cardiff: University of Wales Press.

Schmidt, Moritz & Latte, Kurt (eds). 1965–66. *Hesychius Alexandrinus. Lexicon*, 4 Vols. Amsterdam: Hakkert.

Williams, J. E. Caerwyn, Lynch, Peredur I. & Gruffydd, R. Geraint. 1994. *Gwaith Meilyr Brydydd a'i Ddisgynyddion*. Cardiff: University of Wales Press.

Williams, Ifor (ed.). 1961. *Canu Aneirin*, 2nd edn. Cardiff: University of Wales Press.

Williams, Ifor (ed.). 1990. *Canu Taliesin o Lyfr Taliesin*, 3rd edn. Cardiff: University of Wales Press.

References

Baffa, Giuseppe. 2009. *Dizionario arbëresh di Santa Sofia d'Epiro*. Rende: Università della Calabria, Centro Editoriale e Librario.

Çabej, Eqrem. 1969. Keltisch-Albanische Isoglossen. In (1969). *Studi Linguistici in Onore di V. Pisani*, AA. VV. (eds), 167-186. Brescia: Paideia.

Çabej, Eqrem. 1973. Emri i Dardanisë dhe Izoglosat Shqiptaro-Kelte. *Studime Filologjike* 3: 55-64.

Çabej, Eqrem. 1976 *Studime Etimologjike në Fushë të Shqipes*, Vol. 2. Tirana: Academy of Sciences of Albania.

Çabej, Eqrem. 1982. *Studime Etimologjike në Fushë të Shqipes*, Vol. 1, 2nd edn. Tirana: Academy of Sciences of Albania.

Çabej, Eqrem. 1987. *Studime Etimologjike në Fushë të Shqipes*, Vol. 3, 2nd edn. Tirana: Academy of Sciences of Albania.

Çabej, Eqrem. 1996. *Studime Etimologjike në Fushë të Shqipes*, Vol. 4, 2nd edn. Tirana Academy of Sciences of Albania.

Campanile, Enrico. 1974. *Profilo Etimologico del Cornico Antico*. Pisa: Pacini.

Delamarre, Xavier. 2001. *Dictionnaire de la Langue Gauloise*. Paris: Errance.

Demiraj, Bardhyl. 1997. *Albanische Etymologien*. Amsterdam: Rodopi.

Ernault, Emile 1895. *Glossaire Moyen-Breton*, 2nd edn. Paris: Bouillon.

Falileyev, Alexander. 2000. *Etymological Glossary of Old Welsh*. Tübingen: Niemeyer.
https://doi.org/10.1515/9783110952643

Fleuriot, Léon. 1964. *Dictionnaire des gloses en vieux breton*. Paris: Klincksieck.

Fleuriot, Léon. 1974. Sur quelques textes gaulois. *Études Celtiques* 14(1): 57-66.

Gamkrelidze, Thomas V. & Ivanov, Vjačeslav I. 1995. *Indo-European and Indo-Europeans, Part I: The Text*. Berlin: Mouton De Gruyter. https://doi.org/10.1515/9783110815030

Giordano, Emanuele. 1963. *Fjalor i Arbëreshvet t'Italisë*. Bari: Edizioni Paoline.

Hamp, Eric P. 1966. The Position of Albanian. In Ancient Indo-European Dialects. Proceedings of the Conference on IE Linguistics Held at the University of California, Los Angeles, April 25–27*, 1963*, Henrik Birnbaum & Jaan Puhvel (eds). Berkeley CA: The University of California Press.

Hamp, Eric P. 1972. Albanian. *Current Trends in Linguistics* 9: 1625-1698.

Hamp, Eric P. 1996. Sull'origine settentrionale dell'Albanese. *Quaderni del Dipartimento di Linguistica, Università della Calabria* 15: 109-113.

Henry, Victor. 1900. *Lexique étymologique des termes les plus usuels du Breton moderne*. Rennes: Plihon & Hervé.

Huld, Martin E. 1984. *Basic Albanian Etymologies*. Los Angeles CA: Slavica.

Jokl, Norbert. 1911. *Studien zur albanesischen Etymologie und Wortbildung*. Vienna: Akademie der Wissenschaften.

Jokl, Norbert. 1916. Beiträge zur albanesischen Grammatik. *Indogermanische Forschungen* 37: 98-164.

Jokl, Norbert. 1923. *Linguistisch-kulturhistorische Untersuchungen aus dem Bereiche des Albanischen*. Berlin: De Gruyter. https://doi.org/10.1515/9783111663692

Lambert, Pierre-Yver. 2003. *La langue gauloise*. Paris: Errance.

Loth, Joseph-Marie. 1884. *Vocabulaire vieux-breton*. Paris: Vieweg.

Loth, Joseph-Marie. 1924. Notes étymologiques et lexicographiques. *Revue Celtique* 41: 204-237.

Mallory, Joseph P. & Adams, Douglas Q. 1997. *Encyclopedia of Indo-European Culture*. Chicago IL: Fitzroy Dearborn.

Mallory, Joseph P. 1999. *In Search of the Indo-Europeans. Language, Archaeology and Myth*, 2nd edn. London: Thames & Hudson.

Meyer, Gustav. 1891. *Etymologisches Wörterbuch der albanesischen Sprache*. Strasbourg: Trübner. https://doi.org/10.1515/9783111562261

Niedermann, Max. 1916. *Marcelli de Medicamentis Liber*, 2nd edn. Leipzig: Teubner.

Orel, Vladimir. 1998. *Albanian Etymological Dictionary*. Leiden: Brill.

Pignoli, Maria Luisa & Tartaglione, Guido. 2007. *Dizionario Albanese Molisano (Parlata di Portocannone e Ururi)*. Rende: University of Calabria.

Pisani, Vittore. 1959. *Saggi di linguistica storica. Scritti scelti*. Rosenberg & Sellier: Turin.

Pisani, Vittore. 1959. L'albanais et les autres langues indo-européennes. *In Pisani*, 96-114.

Pisani, Vittore. 1959. *Lexikalische Beziehungen des Albanesischen zu den anderen indo-germanischen Sprachen*. In Pisani, 115-136.

Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*, 2 Vols. Bern: Francke.

Porzig, Walter. 1974. *Die Gliederung des indogermanischen Sprachgebiets*, 2nd edn. Heidelberg: Winter.

Royal Irish Academy. 1990. *Dictionary of the Irish Language* 2nd edn. Dublin: Royal Irish Academy.

Schrijver, Peter. 1995. *Studies in British Celtic Historical Phonology*. Amsterdam: Rodopi.

Sciambra, Matteo. 1964. *La "Dottrina Cristiana Albanese" di Luca Matranga*. Rome: Vatican Apostolic Library.

Scutari, Pasquale. 1991. *Spigolature lessicali sulla parlata degli Arbëreshë di S. Costantino Albanese*. Scanzano Ionico: Lanfranco.

Scutari, Pasquale. 2002. *Il lessico della parlata arbëreshe di San Costantino Albanese*. Rende: University of Calabria, Centro Editoriale e Librario.

Skok, Peter. 1971–1973. *Etimologijski Rječnik Hrvatskoga ili Srpskoga Jezika*, 4 Vols. Zagreb: Croatian Academy of Sciences and Arts.

Starostin, Sergei, Dybo, Anna & Mudrak, Oleg. 2003. *Etymological Dictionary of the Altaic Languages*, 3 Vols. Leiden: Brill.

Stokes, Whitley. 1860. The Welsh glosses and verses in the Cambridge Codex of Juvencus. *Transactions of the Philosophical Society*, 204-232.

Stokes, Whitley. 1865. Die glossen und verse in dem codex des Juvencus zu Cambridge. *Beiträge zur Vergleichenden Sprachforschungen* 4: 385-423.

Stokes, Whitley & Bezzenberger, Adalbert. 1894. *Urkeltischer Sprachschatz*. Göttingen: Vandenhoek & Ruprecht.

Stokes, Whitley & Strachan, John. 1901. *Thesaurus Palaeohibernicus*, 2 Vols. Cambridge: CUP. Reprint: Dublin 1984.

Thomas, Antoine. 1916. Nouvelles variétés étymologiques: Ancien Français *Blaireau*, Provençal *Blar*, Ancien Français. *Bler Romania* 44: 332-333.

Thomas, Richard James et al. (eds) 1950–2000. *Geiriadur Prifysgol Cymru. A dictionary of the Welsh language*. Cardiff: University of Wales Press.

Trumper, John B. 1999. The position of Albanian: Some general lexical problems connected with phylogenesis. In *Studi in onore di Luigi Marlekaj*, Matteo Mandalà (ed.), 531-564. Bari: Adriatica.

Trumper, John B. 2002. Alcune istituzioni socio-poetico-storiche e le relazioni linguistiche tra l'albanese e le lingue indoeuropee dell'Europa centro-settentrionale. In *Studi in onore di Antonino Guzzetta*, Matteo Mandalà & Francesca Di Miceli (eds), 417-442. Palermo: Helix Media.

Trumper, John B. 2016. *Geostoria Linguistica della Calabria*, Vol. I. Rome: Aracne.

Vendryes, Josep, Bachellery, Édouard & Lambert, Pierre-Yves. 1954–1996. *Lexique étymologique de l'Irlandais ancien*. Dublin: Dublin Institute for Advanced Studies.

Walther Von Wartburg. 1929. *Französisches etymologisches Wörterbuch (FEW)*, 25 Vols. Bonn: Klopp.

Language index

Subject index

A

ablative   226

accusative   8, 58, 63, 161, 164, 166, 166n3, 167, 168, 169, 169n8, 169n9, 170, 170n10, 173, 178, 203, 205, 205n3, 206, 210, 211, 223, 226, 227, 227n6, 230, 235n12, 237, 290n5

acquisition, L1/L2   14, 58, 59, 356, 359, 360, 365, 366, 371

adverb   93n22, 94, 143, 172n13, 286, 287, 289, 290, 370

affixes   95, 224, 225n2

agreement   3, 12, 17, 19, 23, 34n5, 42, 42n8, 43, 43n10, 46, 50, 52, 53, 87, 97, 98, 101, 104n3, 105, 106, 107, 108, 109, 116, 125, 144, 145, 146, 147, 148, 178, 179, 181, 182, 183, 184, 185, 189, 238, 241, 242, 243, 244, 245, 246, 247, 249, 250, 253, 254, 259, 260, 261, 262, 263, 264, 266, 267, 279, 280, 366, 367, 368, 370, 371, 378

allomorphy   8, 111, 112, 114, 117, 195, 208, 210, 211, 212, 215, 221, 312

allophony   311, 363

areal diffusion   379, 385

Aspect   11, 12, 7885, 88, 88n11, 89n13, 90, 90n15, 93, 94, 210, 258, 274, 345, 361, 375

B

bogus clusters   308, 319

branching onsets   308

C

cartography   16, 70, 71, 72, 149, 150, 159, 160, 173, 362

causatives   10, 19, 68n11, 71, 161, 169, 172, 174, 265, 266

cleft interrogatives   121, 123, 124, 125

clefts   13, 121, 122, 122n2, 123, 124, 125, 126, 127, 128, 129, 130, 131

clitic climbing   146n8, 148, 166n4, 167, 171, 171n11, 172, 192, 265, 266

cliticization   212

cluster syllabification   307, 312, 316

coda-onset sequences   308, 316

Comment   151, 152

complex onsets   309, 312, 313

consonant clusters   215, 220, 307, 309, 318

contact-induced change   242, 245, 247n5, 254

copulas   12, 85, 86, 87, 89n12, 91, 94, 95

corpora   6, 223, 227n6, 228, 229n8, 234n11, 237, 239, 337, 340

Creoles   370, 371, 373

criterial heads   150, 151

D

dative   6, 7, 11, 12, 24, 68, 68n11, 69, 70, 113, 164, 165, 166, 166n3, 167, 168, 169, 169n8, 169n9, 170, 170n10, 203, 205, 205n3, 206, 210, 211, 238

defective intervention   161, 167, 167n5, 167n6, 168, 171, 173, 174

deletion   5, 6, 34n5, 128, 180, 191, 199, 279

determiner   34, 51, 101, 272, 375

D (continued)

Developmental Dyslexia (DD)   15, 367, 368, 369, 375, 376, 378

Differential Object Marking (DOM)   8, 10, 13, 20, 57, 58, 59, 59n3, 60, 61, 61n5, 62, 63, 65, 69, 70, 72, 266, 278

diglossia   371

direct object   57, 58, 59, 59n2, 60, 61, 62, 68, 70, 73, 75, 164, 165, 166, 169, 243, 261, 378

dyslexia   15, 365, 366, 367, 368, 373

E

ECM constructions   161, 171

electroencephalography (EEG)   353, 354, 355, 361, 362

Element Theory   297, 298, 299, 300, 303, 306

Empty Category Principle (ECP)   72, 188, 189

empty nucleus   308

enclisis   7, 12, 18, 21, 111, 112, 114, 120, 124, 144, 146n8, 208, 209, 210, 211, 212, 257, 258, 259, 264, 265, 294

enclitics   113, 144, 195, 198, 202, 204n2, 205, 211, 212, 290n5

epenthesis   8, 215, 216, 217, 218, 219, 220, 221, 302, 310, 313, 314, 316

EPP feature   67, 67n10

ERMFs   354

ERPs   100, 108, 351, 354, 362

euphemisms   338, 343, 346

external sandhi   284

F

faire-infinitives   161, 162

feature strength   98, 99, 106

Current theoretical approaches to language devote great attention to macro- and micro-variation and show an ever-increasing interest in minority languages. In this respect, few empirical domains are as rich and lively as the Italo-Romance languages, which together with Albanian were the main research domain of Leonardo M. Savoia. The volume covers areas as different as phonology, morphology, syntax and the lexicon. A broad range of Romance languages is considered, as well as Albanian, Greek and Hungarian, shedding new light on many classical topics. The first section focuses on morphosyntax, both in the narrow sense and with regard to its interfaces. The second section focuses on clitics and pronouns. The third section deals with a number of issues in phonology and syntax-phonology interface. The last section turns the reader's attention beyond formal linguistics itself and examines variation in the light of neurosciences, pathology, historical linguistics and political discourse.

John Benjamins Publishing Company