VII-2021

Quaderni di Linguistica e Studi Orientali- QULSO

Working papers in Linguistics and Oriental Studies

Università degli Studi di Firenze

Dipartimento di Formazione, Lingue, Intercultura, Letterature e Psicologia

Biblioteca di Studi di Filologia Moderna: Collana, Riviste e Laboratorio

Quaderni di Linguistica e Studi Orientali *Working Papers in Linguistics and Oriental Studies*

7

Editors

M. Rita Manzini

Leonardo M. Savoia

FIRENZE UNIVERSITY PRESS
2021

Quaderni di Linguistica e Studi Orientali / Working Papers in Linguistics and Oriental Studies - -
n. 7, 2021
ISSN 2421-7220
ISBN 978-88-5518-382-6
DOI: http://dx.doi.org/10.13128/QULSO-2421-7220-7

Direttore Responsabile: Arianna Antonielli

The online digital edition of the journal is published in Open Access on http://oaj.fupress.net/index.php/bsfm-qulso

The products of the Publishing Committee of Biblioteca di Studi di Filologia Moderna: Collana, Riviste e Laboratorio (<https://www.forlilpsi.unifi.it/vp-289-laboratorio-editoriale-open-access-ricerca-formazione-e-produzione.html>) are published with financial support from the Department of Education, Languages, Intercultures, Literatures and Psychology of the University of Florence, and in accordance with the agreement, dated February 10th 2009 (updated February 19th 2015), between the Department, the Open Access Publishing Workshop and Firenze University Press. The Workshop promotes the development of OA publishing and its application in teaching and career advice for undergraduates, graduates, and PhD students in the area of foreign languages and literatures, and of social studies, as well as providing training and planning services. The Workshop's publishing team are responsible for the editorial workflow of all the volumes and journals published in the Biblioteca di Studi di Filologia Moderna series. QULSO employs the double-blind peer review process. For further information please visit the journal homepage (<https://oaj.fupress.net/index.php/bsfm-qulso>).

Editing and layout by LabOA: A. Antonielli (managing editor), S. Crocchi, D. Cappelli, L. Dal Pozzo, F. Salvadori (copy editors), S. Bucalossi, R. De Benedetto, V. Nelli, G. Patacchini, G. Rametti, E. Zoi (interns).

Cover: Salomè H. Varje

We gratefully record that this issue has been co-financed by the Department of Humanities of the University of Florence.

Via Cittadella, 7, 50144 Firenze, Italy
www.fupress.com

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: (2021) Indice. *Qulso* 7: pp. 5-6. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-11999

Indice

Morfosintassi / Morphosyntax

Fonologia / Phonology

Semantica / Semantics

Glottodidattica / Language Teaching

Comunicazione / Communication

Morfosintassi

Morphosyntax

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: L. Badan, P. Lorusso (2021) Subject clitics in nominal copular sentences: some insight from a North Eastern variety of Italian. *Qulso* 7: pp. 9-32. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12000

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Subject clitics in nominal copular sentences: some insight from a North Eastern variety of Italian

Linda Badan and Paolo Lorusso
Ghent University (<linda.badan@ugent.be>),
Università degli Studi di Firenze (<paolo.lorusso@unifi.it>)

Abstract:

In this paper we describe the distribution of subject clitics in nominal copular constructions in the Veneto variety of Este (Padova province). In nominal copular sentences in a *pro-drop* language like Italian, the copula always agrees with the subject of the small clause both in canonical (preverbal NP subject) and in inverse (postverbal NP subject) structures. Since Veneto varieties are partial *pro-drop* languages, on the one side, the subject NP always agrees with the copula as in Italian. On the other side, the inflectional paradigm of both verbs and subject clitics interacts with the informational structure of the sentence. Namely, the allowance for subject proclitics and overt postverbal subjects in inverse copular sentences (which is not found in other constructions involving postverbal subjects) allows us to account for the preverbal position of the predicative DP in terms of informational structure: the preverbal predicative DP is a topic intensional element that instantiates the description of the subject of the sentence. The predicative DP, as in Italian, can cliticize onto an invariant pro-predicative clitic: the resulting clustering with the subject clearly confirms the reduced structure involved by copular constructions (as in Moro 1997). The present descriptive considerations allow us to update the analysis of Moro (1997). We characterize inverse copular sentences as involving some restriction on the topic-comment structure, showing a clear morphological counterpart (the inflected subject proclitics) in the operation of AGREE in a partial *pro-drop* language like the variety of Este (Padova province).

Keywords: *Agreement, Nominal Copular Constructions, Partial Pro-Drop, Subject Clitics, Veneto Variety*

1. Introduction

In this paper the distribution of subject clitics in copular sentences is described in the Veneto variety of Este (Padova province). The variety of Este, as other varieties with subject clitics, is a partial *pro-drop* language, since the null subject and

the inflected subject clitics alternate depending on argument structure, the salience of person within the inflectional paradigm (Manzini and Savoia 2005), and the informational structure of the sentence. Furthermore, as in other Veneto varieties (but not only), two clitics coexist (Benincà 1994; Poletto 2000; Manzini and Savoia 2005): a lower inflected subject clitic (in both proclisis and enclisis) and a higher uninflected one. The invariant higher clitic does not seem to be linked to the inflectional field (while inflected "lower" clitics are found to express D(efinite) denotation), but it is linked with scope discourse-semantics effects (Benincà 1994), or modality (Poletto 2000) or intensional semantics (Manzini and Savoia 2005).

In this paper, we describe the distribution of subject proclitics in nominal copular sentences. We show that both canonical (1a) and inverse copular constructions (1b) (as defined in the unified theory of Copular Construction proposed by Moro 1997) have different informational properties, arguing that inverse sentences are strictly linked to a marked informational structure.

(1) a. I ragazzi sono la causa del problema
the boys be.3PL the cause of-the problem
'The boys are the cause of the problem'

b. La causa del problema sono i ragazzi
the cause of-the problem be.3PL the boys
'The cause of the problem is the boys'

We argue that the marked information structure of inverse copular sentences introducing a postverbal subject is similar to the one of the constructions involving dislocated NPs and differs from the information structure of postverbal subjects with lexical verbs. Through the preverbal NP, inverse copular structures introduce a non-extensional meaning with a postverbal referential NP that encodes new information. This fact is confirmed by the agreement relation: the verb probes the postverbal referential NP (1b). The agreement is available only with one NP: as in Italian, the predicative NP can cliticize onto a non-inflected element which is syncretic with object clitics, but which does not show any overt agreement features.

The paper is organized as follows: in section 2 we present the data about the distribution of the clitics in the variety of Este; in section 3 we illustrate the properties of the agreement, while in section 4 we describe the distribution of clitics within copular sentences. Section 5 is devoted to the analysis of one of the other diagnostics individuated by Moro (1997) to distinguish between canonical and inverse copular sentences, namely the presence of a pro-predicative invariant clitic to refer to the predicate NP in canonical sentences. In section 6 we resume the syntactic relevant facts that emerge by our description, and in section 7 we present our concluding remarks.

2. The subject clitics in Veneto varieties: paradigm and syntactic distribution

In this section we briefly illustrate the system of clitic subjects in a Veneto variety spoken in the Padova province. We focus on the variety spoken in Este and the immediately surrounding municipalities.[1] As for their distribution, the clitic subjects in the variety examined here correspond to the system illustrated by Renzi and Vanelli (1983), Poletto (1993, 2000), Benincà

[1] Este is around 40 km far from Padova. The informants we have consulted are mainly from Este, but also from Ospedaletto, Baone, Lozzo Atestino, Carceri.

(1994), for Paduan: they can be enclitic and proclitic with respect to the verb, depending on the syntactic context in which they are realized. Both enclitics and proclitics agree in person and number with the verb, but while the enclitics present a full paradigm for all the singular and plural persons, the proclitics lack of the first person singular and plural and of the second person singular.[2] Both enclitics and proclitics codify the gender in their third person singular. In Tab.1 we illustrate the whole paradigm of cliticization in the Este variety.

	singular				**plural**			
	1	**2**	**3**		**1**	**2**	**3**	
			M	**F**			**M**	**F**
proclitics	-	te	el	ea	-	-	i	e
enclitics	-i	-to	-(e)o	-(e)a	-i	-o	-i	-e

Table 1 – Paradigm of clitics in the Este variety

The realization and the distribution of the subject clitics illustrated in Tab. 1 are determined by both syntactic and pragmatic factors. The proclitic subject is obligatorily realized in all (allowed) persons and numbers when the tonic subject pronoun or an overt subject DP is absent (2b), while it is optional in co-occurrence with a subject pronoun (2c) or an overt subject DP (2a) when it is topicalized (we illustrate a possible context in (2d; Poletto 2000; Benincà 2004; see also Raposo and Uriagereka 1996).

(2) a. Me mama (ea) vien domàn
my mum CL.3SG.F come.3SG tomorrow
'My mum comes tomorrow'

b. *(Ea) vièn domàn
CL.3SG.F come.3SG tomorrow
'She comes tomorrow'

c. Eo (el) vièn domàn
he CL.3SG.M come.3SG tomorrow

d. A: E to mama?
and your mum
'And your mum?'
B: Me mama, (ea) vièn domàn
my mum CL.3SG.F come.3SG tomorrow
'My mum comes tomorrow'

The optionality of subject clitics is not found when the 2nd person singular is involved: in this case the clitic is obligatory in co-occurrence with the second singular of the tonic pronoun:[3]

[2] As for proclitics, in the paradigm of the Este variety there are not subject clitics for the first person singular and plural and for the second person plural. Moreover, across the paradigm of the proclitics, no syncretic forms are found showing similarity with the paradigm of franco provenzal variety of Sarre, among others, described by Manzini and Savoia (2005) and Manzini (2015).

[3] As for the status of 2nd person in the paradigm of subject clitics, the traditional view of Renzi and Vanelli

(3) a. Ti te vièn domàn
you.2SG CL.2SG come.2SG tomorrow
b. *Ti vièn domàn
you.2SG come.2SG tomorrow
c. *Te vièn domàn
CL.2SG come.2SG tomorrow
'You come tomorrow'

Instead, the presence of the subject clitic is always excluded if (i) it is co-indexed with a wh-trace in interrogatives (4a-b) and in relatives (4c); (iii) if co-indexed with the clefted element in cleft sentences (4d); (iv) within sentences with atmospheric verbs (5); (v) when the subject is focalized (see (6a) with a contrastive focus and (6b) with an informational focus).

(4) a. Non so chi che (*el) vièn
NEG know.1SG who that CL.3SG.M come.3SG[4]
'I don't know who comes'
b. Chi (*el) vièn?
who CL.3SG.M come.3SG
'Who comes?'
c. El tozo che (*el) vièn domàn
the boy that CL.3SG.M come.3SG tomorrow
'The boy that comes tomorrow'
d. Ze Giorgio che (*el) vièn domàn
be.3SG Giorgio that CL.3SG.M come.3SG tomorrow
'It's Giorgio who comes tomorrow'

(5) *El piove
CL.3SG rain.3SG
'It rains'

(6) a. GIORGIO (*el) vièn domàn (no Carlo)
Giorgio CL.3SG.M come.3SG tomorrow NEG Carlo
'GIORGIO comes tomorrow (not Carlo)'
b. A: Cossa ze capità?
'What happened?'
B: (*Ea) ze rivà ea tosa
CL.3SG.F be.3SG arrived the girl
'The girl is arrived'

Moreover, notice that the proclitic subject is never allowed with postverbal subject in declarative sentences as in (7). This is particularly relevant for our analysis, since in inverse copular constructions subject clitics may instead appear with postverbal DP (see section 4).

(1983) is in an implicational scale for which 2nd person is more marked. The same implication is found also in Cardinaletti and Repetti (2008). A more recent approach such as the one of Manzini (2015) is derived by a parametric option linked to the relevance of first person (see section 2.2).

[4] Examples adapted from Benincà (1994: 18, ex. 3).

(7)	(*Ea)	ze	rivà	to	mama.
	CL.3SG.F	is	arrived	your	mum
	'Your mother is arrived'				

However, notice that the subject clitic obligatorily appears when the subject is a right dislocated topic (8).

(8)	(El_i)	vièn	domàn,	$Gianni_i$
	CL.3SG.M	come.3SG	tomorrow	Gianni
	'He comes tomorrow, Gianni'			

As for the enclitic forms, they are used in interrogative clauses as in the following:

(9)	Ga-(e)o	za	magnà	fora	tuto?
	have-CL.3S.M	already	eaten	out	everything
	'Did he already eat everything?'				

However, in this paper we will not address the issue of enclitics in interrogative inversion. The main topic of the present work focuses on the analysis of nominal copular sentences, where enclisis seems to be problematic due to the reduced inflectional layer found with copulas (Moro 1997: 73). However, it is worth pointing out that while subject enclitics are never found in declarative context (10a-b), they can be found in interrogative inversion of canonical copular sentences (10c), but not in inverse copular sentences (10d).

(10)	a.	*Ea	causa	dea	baruffa	ze-i		i	tosi
		the	cause	of-the	fight	be.3PL-CL.3PL.M		the	boys
		'The cause of the fight are the boys'							
	b.	*I	tosi	ze-i		ea	causa	dea	baruffa
		the	boys	be.3PL-CL.3PL.M		the	cause	of-the	fight
		'The boys are the cause of the fight'							
	c.	Ze-i		i	tosi	ea	causa	dea	baruffa?
		be.3PL-CL.3PL.M		the	boys	the	cause	of-the	fight
	d.	*Ze-i		ea	causa	dea	baruffa	i	tosi?
		be.3PL-CL.3PL.M		the	cause	of-the	fight	the	boys
		'Are the boys the cause of the fight?'							

However, for a wide discussion on the differences between proclitics and enclitics see Calabrese (2011), Cardinaletti and Repetti (2008), or Manzini and Savoia (2005). See also Manzini (2015) who proposes that enclisis, in fact, is found under the scope of non-veridical operator (Giannakidou 1998) or a general raise to a functional projection in the CP domain.

2.1 The invariant "clitic" a

As in Paduan and other Veneto varieties, also the Este variety presents an invariant (subject) clitic consisting in the vowel *a*. Benincà (1994) and Poletto (1993, 2000) have examined *a* in detail concluding that actually *a* is not a "true" subject clitic. Indeed, *a* in Paduan is optionally used mainly, but not exclusively, for the first singular and plural persons and for the second

plural persons, and it appears in specific pragmatic conditions: it is strictly linked to the prosodic contour of the sentence expressing emphasis or surprise, indicating that the meaning of the whole clause is a rheme, a new information. Additionally, to its invariant morphology and its specific pragmatic role, *a* presents also a syntactic distribution that is crucially different from the distribution of subject clitics.

(I) *a* can appear with meteorological (11a) or impersonal verbs (11b). The appearance with meteorological verbs seems to confirm the analysis of *a* as an "expletive"[5] since the clitic does not refer to any overt argument. However, we follow Manzini and Savoia (2005, I: 183) arguing that the invariant clitic *a* does not refer to any referential individual or to its (D)efinite properties.

(11) a. A piove!
CL rain.3SG
'It rains!'
b. A bisogna corare!
CL need.IMPRS run
'One has to run!'

(II) *a* is the only clitic that (optionally) appears in preverbal position when the subject is postponed (12). *a* is invariant and does not show any overt inflectional marking agreeing with the postverbal subject. The invariant *a* is also found in some other varieties (Manzini and Savoia 2005) in constructions where the agreement is missing between the verb and the postverbal subject as illustrated in (13) for the variety of Revere.

(12) a. (A) riva Giorgio
CL arrive.3SG Giorgio
'Giorgio arrives'
b. (A) vago mi
CL come I
'I go'

(13) A parlɑ i ɔm *Revere*
CL speak.3PL the men
'The men speak'

(Manzini and Savoia 2005, I: 300)

(III) In the case of the second person singular, i.e. the case with the subject clitic *te* always obligatory, *a* can co-occur with it, but it cannot substitute it (14).

[5] Manzini and Savoia, (2005), following Chomsky (1981), argue that meteorological verbs do not lack of argument structure, but they are associated with thematic properties ("aspectual" in Manzini and Savoia's terms) which can be fulfilled by quasi-arguments. Recall that in Italian it is possible to say 'pro *piove manna' (= it rains manna)*. We will not enter into the relevant discussion, but, for the purpose of the present work, we consider the quasi-argument as non-arguments.

(14) a. A te parli sempre (ti)
CL CL.2SG speak.2SG always you.2SG
b. *A parli sempre ti
CL speak.2SG always you.2SG
c. Te parli sempre ti
CL.2SG speak.2SG always you.2SG
'You are the one who always speaks'

In this case, as argued by Manzini and Savoia (2005), the 2nd person has a special status, since due to his "relevance" which is inferior to the 1st person (which is usually not found in the clitic paradigm), it is commonly lexicalized with an obligatory 2nd person clitic.

The example (14a) also shows that the clitic *a* can appear with a postverbal 2nd person tonic pronoun, but never with a preverbal tonic pronoun (15), which is in a topic-like position in CP. So, the general status of the expletive-like and invariant *a* is the one of an element which deals with the universe of CP, which, in Manzini and Savoia's (2005, I: 141) terms, deals with the intensional meaning of the sentence.

(15) *Ti a te parli sempre
You.2SG CL CL.2SG speak.2SG always
'You always speak'

(IV) a can co-occur with a subject clitic, always on its left:

(16) A'l parte domàn!
CL-CL.3SG leave.3SG tomorrow
'He leaves tomorrow!'

(V) Differently from subject clitics, the invariant *a* always precedes the negation (17), showing that its position is higher than NegP, which is commonly assumed to be in a position within the inflectional layer. This also confirms the C position of the invariant *a* clitic.

(17) a. A no te parli mai
CL NEG CL.2SG speak.2SG never
b. *Te no a parli mai
CL.2SG NEG CL speak.2SG never
'You never speak'

(VI) a can appear in yes/no questions. Differently from subject clitics that trigger inversion appearing as the proclitic forms (see Table 1), the invariant *a* always remains in sentence initial position (18). Yes/no questions in fact, do not imply an overt movement to C of any element of the VP or IP.

(18) A ve-to via?
CL go.2SG-CL.2SG away
'Do you go away?'

(VII) However, *a* is incompatible with a wh-element: that is, *a* may compete with *wh* for a position within the CP field.

(19) *A dove si-to 'ndà?
CL where be.2SG-CL.2SG gone
'Where did you go?'

(VIII) Similarly, *a* cannot occur with a left dislocation (20) or with a contrastive focus in sentence initial position (21).

(20) a. *Gianni, a eo gò visto 'ndar via.
Gianni CL CL.3SG.M have.1SG seen go.INF away
b. *A Gianni, eo gò visto 'ndar via.
CL Gianni CL.3SG.M have.1SG seen go.INF away
'Gianni, I have seen it going away'

(21) a. *GIANNI a ze 'nda via (no Piero)
Gianni CL be.3SG gone away NEG Piero
b. *A GIANNI ze 'nda via (no Piero)
CL Gianni be.3SG gone away NEG Piero
'GIANNI, he's gone away (not Piero)'

On the basis of the analysis of *a*'s syntactic behaviour and pragmatic interpretation, both Benincà (1994) and Poletto (2000) argue, even if in different terms, that *a* is strictly linked to the informational structure of the clause, and it occupies a structural position higher than the subject. Benincà calls such a position TOP, while Poletto defines it as a Modality projection above AgrTP, where the subject is realized. Manzini and Savoia (2005, I: 141) do not enter into details about the properties of the C position where the *a* clitic is realized. However, they find many varieties in which the *a* is also syncretic with the 3rd person form, therefore they propose that *a* can lexicalize the D properties. The D properties can be lexicalized either within the inflectional layer, where the D properties denote an individual, or within the C domain, where the D properties are individuated as part of the intensional meaning of the sentence, i.e. they refer to a set of individuals but not a concrete individual. When they refer to a concrete individual, *a* can coexist with inflected lexical subject clitics. Furthermore, *a* can be used as an expletive to satisfy the EPP principle (Extended Projection Principle) for which each sentence must contain a noun phrase or determiner phrase in the subject position. We will discuss and develop these proposals in section 3 and 4, to illustrate our analysis of inverse copular sentences.

3. Agreement in (partial) pro-drop languages

Subject clitics in Northern Italian varieties have been the object of a number of studies. Rizzi (1986), on par with Brandi and Cordin (1989), analyze subject clitics (in Trentino variety) as INFL heads. Poletto (2000) classifies subject clitics in two main groups, agreement elements or complementizer items, depending on their features. Roberts (2010) proposes that subject clitics are heads that incorporate with the head T, then in his book (2014) he develops the idea of conceiving subject clitics as inflections, consisting in bundles of features merged together with the verb in the head T.

While the accounts mentioned above rely on a movement or a partial movement analyses, other scholars argue for a base generation analysis. For instance, Sportiche (1992, 1996) introduces a clitic phrase within the inflectional layer, while Manzini and Savoia (2005) recognize

an inflectional string for clitic placement both in the INFL and in the C layer: the former being related to the individuation of personal reference (and showing overt inflectional morphology) and the latter being related to the attribution of intensional meaning associated with the sentence.

The analysis of subject clitics is inserted within a wide discussion on the *pro-drop* (null subject) parameter.[6] In the literature, languages are defined as: (i) Non-*pro-drop* languages, which never allow subject omission like French or English; (i) full *pro-drop* languages as for instance Italian, Spanish, Greek, where the referential subject can be null and the null subjects can be referential or non-referential (expletive). Languages with a pronominal referential INFL can license referential null subjects; (ii) Radical *pro-drop* languages, where instead both subject and object can remain silent and generally they do not have any verbal inflection like Chinese (see Huang 1984); (iii) Expletive null subject languages, where only the expletive subject can be silent, while full referential subjects must always overtly realized, as in Dutch, or several Creole languages (Nicolis 2008); (iv) Partial *pro-drop* languages where the null subject is regulated by syntactic conditions: some of them do not have full referential silent pronouns, but have only null indefinite pronouns, or null arbitrary pronouns. Finnish, Marathi, Assamese, as well as Brazilian Portuguese are classified as partial null subject languages (see Holmberg 2010; Holmberg and Sheehan 2010; Biberauer *et al.* 2010).

As all the varieties which involve subject clitics are partial *pro-drop* languages, also in the Este variety the distribution of null subject across the inflectional paradigm may involve: (i) the lack of subject clitics with the first singular and plural person and the second plural person; (ii) no 3rd person inflected subject clitic with meteorological and impersonal verbs (although the uninflected *a* is allowed in such configuration); (iii) no subject clitic with postverbal subjects (once more the *a* can do the job). Therefore, while in a *pro-drop* language like Italian, the inflection of the finite verb has a role in both identifying the phi-features of the referential subjects and to satisfy the EPP, in a variety of Este (as in other partial *pro-drop* languages) we have a finer graded identification of the phi-features of the referential subjects and the EPP requirements. In terms of Manzini and Savoia (2007), the EPP property corresponds to a D(efiniteness) closure requirement: the subjects DP or the finite verb morphology have the denotational content D(efiniteness). If we use the D(efinitiness) feature, we can define the *pro-drop* parameter as how different languages realize this feature (Manzini and Savoia 2007). The D position of the sentential INFL domain can be lexicalized by a specialized head (such as subject clitics the variety of Este), by a full noun phrase (as in English) or by either a specialized head or a full phrase (as in French). By contrast, in a language like Italian, the D position of the sentential I domain is not lexicalized, while the D argument is lexicalized only at the morphological level by the inflection of the finite verb. In terms of the parametric condition on the lexicalization of the D properties, Manzini and Savoia (2007: 45) propose a schematization like in (22). The divide between (a) and (b) in (22) corresponds to the classical divide between null subject languages and non-null subject ones.

(22) Lexicalization of the D properties of the sentential I domain:
- a. i by clitic (e.g. Northern Italian dialects)
  - ii by clitic or noun phrase (e.g. Ladin dialects, French)
  - iii by noun phrase (e.g. English)
- b. no lexicalization (e.g. Italian)

[6] For a recent collectanea of studies on the null subject parameter see Biberauer, Holmberg, Roberts and Sheehan (2010).

In our analysis, the *pro-drop* parameter can be restated in terms of Lexical Parametrization (Manzini and Wexler 1987): the parameter is given depending on how the D features are lexicalized, for instance in the variety of Este they are lexicalized by the subject clitic. However, within this general pattern of lexicalization, there are other lexical and morpho-syntactic features, which are in a sub-set relation to the general *pro-drop* parameter, that influence the distribution of subject clitics. In the variety of Este, the verbal frame, the informational status and the person morphology imply a lack or an overt realization of the subject clitic.

On the basis of the considerations above, we argue that the lexical class may influence the lack of subject clitics. With meteorological verbs, in the variety of Este there is a lack of subject clitics (see (5) above repeated below as (23)). However, as described in Manzini and Savoia (2005), this is a matter of parametric variation across varieties. There are, in fact, varieties in which there is a non-argumental (or "quasi-argumental" as in Chomsky 1981) expletive subject clitics with meteorological verbs. So, in the variety of Este, the subject clitics have to be [+argumental].

(23) (* El) piove *Este*
CL.3SG rain.3SG
'It rains!'

(24) ɛl pjøf. *La Strozza*
CL.3SG rain.3SG
'It rains!'

(Manzini and Savoia 2005: 43)

Similarly, in the Este variety, subject clitics cannot occur with postverbal subjects (25). Once more, while in the variety of Este we argue that when a postverbal subject is present, the subject clitic does not lexicalize the D properties (which are lexicalized just on the inflection on the verb and the full DP), in other varieties, the subject clitics still lexicalize the D properties also in the presence of a postverbal (focal) subject, as in the variety of La Strozza (26).

(25) (*Ea) ze rivà to mama.
CL.3SG.F is arrived your mum
'Your mum has arrived'

(26) De la ɛl dɷrma i stʃɛ:tʃ *La Strozza*
there CL.3SG.M sleeps the children
'The children sleep there'

(Manzini and Savoia 2005: 43)

However, remind that if the postverbal subject is a right dislocated topic, subject clitics are allowed (see (8) repeated here as (27)).

(27) (El_i) vièn domàn, $Gianni_i$
CL.3SG.M come.3SG tomorrow Gianni
'He comes tomorrow, Gianni'

As for the last subset which intervenes in the lexical parametrization in the Este variety, Person has a central role on the distribution of overt/null subject clitics. As Manzini and Savoia (2005,

2011) and Manzini (2012) have suggested, language may differ on the lexicalization of Person[7] through subject clitics: while some languages lexicalize all persons,[8] other languages do not.[9] The main differences are linked, then, to a finer fault line, for example that between "speaker" participant and "hearer" participant. This may result in the externalization of just "speaker" reference or of "hearer" reference. However, while in the enclitic paradigm there is a richer variation on the pattern of lexicalization of Person among the inflectional paradigms of subject clitics (Manzini 2012), in the analysis of proclitic paradigm across varieties there is not such a degree of variation. All varieties have a 2nd person subject clitic, in other words, there is no variety with a null subject proclitic for 2nd person "hearer". As we mentioned above, in the variety of Este, the second person clitic is always obligatory, while there is no 1st person subject proclitic.

(28) a. Ti te vièn domàn
Tu CL.2SG come.2SG tomorrow
b. *Ti vièn domàn
Tu come.2SG tomorrow
c. *Te vièn domàn
CL.2SG come.2SG tomorrow
'You come tomorrow'

This seems to be a parametric option: as proposed by Manzini (2015: 183) the 1st person may not be found due to the "salience of a speaker reference". So, the variety of Este is in a subset relation of the varieties which have proclitic for all persons;[10] this subset is given by the non-lexicalization of the speaker, due to its referential salience. Therefore, the general configuration of agreement we observe for subject clitics is that the D feature can be deleted through agreement, after the subject clitic has been merged in spec,INFL, together with the rest of the phi-features that enter the AGREE relation (Chomsky 2001). The parametric variation among the distribution of subject clitics across varieties can be accounted in terms of lexical parametrization in the lexical element that lexicalizes the D features. Languages vary on when subject clitics lexicalize the D properties, depending on: (i) the [+/-argumental] status of the element bearing the D features, (i.e. in the variety of Este, subject clitics lexicalize only [+argumental]); (ii) the informational properties [+/-focal] of the D, (i.e. in the variety of Este, only [–FOCUS]); (iii) the lexicalization of D or P and in the latter case on the [+/-saliency] of the speaker.

In a partial *pro-drop* language as the variety of Este, the clitic *a* may imply the lexicalization of non-argumental (or quasi-argumental) NP. Furthermore, due to its invariant status, *a* is used to refer to a set of entities intentionally and not as an individual bearing overt nominal agreement features.

Given the considerations illustrated so far on the agreement configuration in partial *pro-drop* languages which lexicalize D properties through subject clitics, in the next section we turn to describe how the reduced agreement configuration of copular sentences might work in these varieties.

[7] Manzini and Savoia (2005 and subsequent work) refer to a categorial splits "speaker" vs. "hearer", P (1st /2nd person) versus D (3rd person).

[8] As in the variety of Barcis (a North-Eastern variety spoken in the area of Pordenone) where the paradigm of subject clitics include not only 3rd persons but also 1st and 2nd person (see Manzini and Savoia 2005 and Manzini 2015).

[9] Although these varieties show a full-fledged subject clitic paradigm for 3rd person, they do not show any subject clitic for 1st and 2nd persons, as in the variety of Pozzaglio (in the area of Cremona).

[10] We refer always to 1st and 2nd persons, since as Manzini and Savoia (2005, 2007, 2011), Harley and Ritter (2002) Bobaliik (2008) emphasize, 3rd person is a non-person.

4. Nominal copular sentences: agreement and clitics

As for nominal copular sentences, Moro (1997) claims that they involve a raising predicate (the copula) that selects a small clause as its complement, and it does not have any particular meaning. The small clause is the place where the predication occurs between two XPs: a subject and a predicate. Following Moro (1997), on the one hand, a canonical copular sentence is a configuration where the subject generated in the small clause raises to the higher position of the copula, while the predicates stays in situ (29a). On the other hand, an inverse copular sentence is a configuration where the predicate generated into the small clause can raise to the higher position of the copula while the subject DP stays in situ (29b):

(29) a. [[Una foto del muro]$_i$ [IP è [SC t_i [la causa della rivolta]]]]
[[A picture of the wall]$_i$ [IP is [SC t_i [the cause of the riot]]]
b. [[La causa della rivolta]$_k$ [IP è [SC [una foto del muro] t_k]]]
[[The cause of the riot]$_k$ [IP is [SC [a picture of the wall] t_k]]]

In *pro-drop* languages like Italian, in nominal copular sentences the verb always agrees with the subject, independently of its pre- or post-verbal position (30a-b). While in non *pro-drop* languages like English, the verb agrees with the subject DP only when it is in preverbal position (31a-b):

(30) a. Le foto sono/ *è la causa
the pictures be.3PL be.3SG the cause
'The pictures are /*is the cause'
b. La causa sono/ *è le foto
the cause be.3PL/ be.3SG the pictures
'The cause are/ *is the pictures'

(31) a. The pictures are/*is the cause
b. The cause *are/is the pictures

Moro (1997) proposes that what triggers agreement on the copula in Italian is not the pre-verbal DP but rather *pro*, which is obligatorily inserted and it indirectly agrees with the subject *in situ*. In Italian, then, *pro* behaves like an anaphor on a par with all inverse subjects and is licensed in a lower pre-verbal position in finite clauses (suggesting that, in Italian, both preverbal and postverbal subjects are dislocated, much in the sense suggested by Cinque 1979, Benincà and Cinque 1993). This happens also in the inverse copular sentences, where the DP predicate raises to a preverbal position, the DP subject is *in situ*, and *pro* refers to the subject of the small clause:

(32) [DP pred [IP *pro*$_i$ copula [SC DP Subj$_i$ ~~DPpred~~]]]

English differs from Italian because no *pro* (or inflection) can lexicalize D properties and the preverbal full DP agrees with the verb. Also in partial *pro-drop* languages as the Este variety, the verb invariably agrees with the subject. The difference with Italian is that the D properties are lexicalized either by *pro* or by the subject clitic. In (33) we illustrate a canonical (33a) and an inverse (33b) copular construction in the Este variety. We notice that the 3rd person is syn-

cretic between subject and plural, so it is not possible to understand whether the agreement is with the subject DP or not.

(33) a. E tose ze ea causa dea barufa
the girls be.3PL the cause of-the fight
'The girls are the cause of the fight'
b. Ea causa dea barufa ze e tose
the cause of-the fight be.3PL the girls
'The cause of the fight are the girls'

In the following examples of canonical sentences, we use tonic pronouns as subjects to show the morphological differences in the verb agreement across the inflectional paradigm. Notice that in the second person singular the presence of the clitic *te*[11] is always mandatory (see section 2). Obviously, the verb agrees with the preverbal subject.

(34) Mi so ea causa dea barufa
I be.1SG the cause of-the fight
'I am the cause of the fight'

(35) Noialtri semo ea causa dea barufa
we be.1PL the cause of-the fight
'We are the cause of the fight'

(36) Voialtri si ea causa dea barufa
you.2PL be.2PL the cause of-the fight
'You are the cause of the fight'

As in canonical copular sentences (as in Italian (30)), in the inverse copular sentences, the verb invariably agrees with the subject that is post-verbal in inverse constructions:

(37) Ea causa dea barufa so mi
the cause of-the fight be.1SG me
'The cause of the fight is me'

For our analysis, we follow a more recent minimalist account (Chomsky 2001) that conceives the formulation of AGREE as a relation between a probe (the verb) and a goal within a minimal domain. On this basis, we can describe the fact that in the Este variety and in Italian, the agreement in inverse copular sentences with the postverbal subject is linked to a probe specification of the raising verb *be* (as proposed by Bejar and Kahnemuyipour 2017 for Iranian and Armenian) which agrees with the element involving both N(ominal) and *deictic* features (38). Since the preverbal NP predicate does not have D features (as also in Moro 1997) but only deictic features (remind that it refers to a set of elements sharing a property), it does not agree with the verb since the agree probe is specified for agreeing with an element involving both deictic and nominal feature.

[11] We will back on this in section 5, where we will be referring to the differences found in the paradigm depending on person and number.

(38) Probe specification for the copula in Inverse copular constructions.

AGR	NP predicate	NP subject
[n]	[d]	**[n]**
[d]		**[d]**

In Este variety we notice different interesting phenomena. Firstly, in the second person singular, the proclitic *te* is always present, as in canonical copular structures. Remind that the 2nd person is a participant and a deictic, and can be the target of the specialized probe.

(39) a. Ti te si ea causa dea barufa
you.2SG CL.2SG be.2SG the cause of-the fight
'You are the cause of the fight'
b. Ea causa dea barufa te si ti
the cause of-the fight CL.2SG be.2SG you.2SG
'The cause of the fight is you'

In the third singular and plural person the presence of the proclitic is always acceptable, but never mandatory:

(40) a. Ea causa dea barufa ea_i ze so $mama_i$
the cause of-the fight CL.3SG.F be.3SG her mum
'The cause of the fight is her mum'
b. Ea causa dea barufa i_k ze i so $fradei_k$
the cause of-the fight CL.3PL.M be.3PL the his brothers
'The cause of the fight is his brothers'

However, in other constructions with lexical verbs, a subject proclitic is never allowed with a postverbal subject.

(41) (*Ea) ze rivà to mama.
CL.3SG.F is arrive.PRT your mum
'(She) is arrived your mum'

This difference suggests that there is a structural distinction between the two types of postverbal subjects. While sentences like (41) can be used to answer to questions like 'what happened?', confirming that they can be interpreted as broad focus, inverse copular sentences can never be broad focus. Roughly, while the preverbal predicate DP has a clear topic reading, the postverbal subject DP may be interpreted as a narrow focus. Specifically, the postverbal DP seems to be in a lower IP criterial position (in terms of Rizzi 2010) encoding a marked informational status (focus) which is not found in the postverbal subject with lexical verbs. Subject clitics can lexicalize D properties when the subject is in a criterial position (postverbal narrow focus or in a right dislocated topic position) and such a position is not available for a full computation of the agreement (as also suggested in Bianchi and Belletti 2016 for postverbal focus), therefore the D properties need to be expressed by the preverbal clitic subject.

Another possibility to account for the difference between copular sentences and sentences involving a lexical verb with postverbal subject is linked to the fact that the inflection of the

lexical verb in (41) parametrically identifies the D properties of the 3rd person subject. In (41) the clitic is obligatory since the raising verb *be* is a mere raising predicate and it does not involve meaning. Furthermore, the 3rd singular person of the verb *be* is syncretic with the plural, so it is a verb with a particular status: it does not have lexical meaning, it does not assign any thematic role, but it is mere inflection (Moro 1997) that cannot lexicalizes the D properties. The D properties are lexicalized by the subject clitic. The same is true for other raising verbs such as *seem* in (42):

(42) Ea causa dea barufa ea me pare so mama
the cause of-the fight CL.3SG.F to-me seem her/his mum
'The cause of the fight to me seems her/his mother'

In sentences involving person and number that lacks subject clitics (in the singular and plural person and in the second person plural), proclitics are not allowed because these forms are lacking in the paradigm. The D properties can only be lexicalized by the inflection of the verb. In the case of the 1st person singular, the inflection is useful to identify the referential individual (i.e. the speaker), due to the contextual relevance of the speaker (Manzini and Savoia 2005; Manzini 2012). However, in the Este variety, in all the structures in which inflected subject clitics are missing, it is possible to have the invariant clitic vowel *a* (discussed in section 2) in preverbal position:

(43) a. Ea causa dea barufa (a) so mi
the cause of-the fight CL be.1SG I
'The cause of the fight is me'

b. Ea causa dea barufa (a) semo noialtri
the cause of-the fight CL be.1PL we
'The cause of the fight is us'

c. Ea causa dea barufa (a) si voialtri
the cause of-the fight CL be.2PL you.2PL
'The cause of the fight is you'

In the case of the second person singular, when the invariant clitic *a* can co-occur with the obligatory proclitic *te*, *a* must precede it:

(44) Ea causa dea barufa (a) te si ti
the cause of-the fight CL CL.2SG be.2SG you.2SG
'The cause of the fight is you'

As for the interpretation of the clitic *a*, due to its invariant status and to its co-occurrence with an inflected subject clitic (obligatory with 2nd person), *a* seems to lexicalize the D properties of the intensional meaning of the set to which include the referential subject(s), and it can be used as an informational/modal marking of the sentence. It does not lexicalize the D features as the inflected subject clitics.

Remind that the inverse copular constructions, sometimes called "specificational copular" constructions (see Bejar and Kahnemuyipur 2017 among others), imply an interpretation for

which the preverbal predicative NP "introduces something like a description" and the subject NP "specifies the entity or entities that fit the description" (Bejar and Kahnemuyipur 2017: 8). The postverbal NP represents new information. This is perfectly compatible with an interpretation for which the preceding preverbal DP predicate is either a topic or a modal operator. This observation allows us to make some generalizations on the position of the preverbal predicate DP. It is important to notice that *a,* due to its intensional meaning, is compatible with the postverbal subject with lexical verbs and can cooccur with subject clitics in inverse copular sentences with a postverbal subject DP. However, *a* is incompatible with topic left dislocation and contrastive preverbal focus (see examples (20)-(21)), since the informational properties of the dislocated element may ban the "intensional" presentation of the sentential subject encoded in the invariant clitic *a.*

Furthermore, in inverse copular sentences, the preverbal predicate DP cannot raise to a contrastive focus position. Although it has a topic reading, it does not work like a topic within a clitic left dislocation because it is not a l-marked argument (in the sense of Cinque 1990), it represents a property (with intensional meaning) and it can only cliticize in an invariant propredicative clitic (no resumptive inflected clitic is available). Next section is devoted to the occurrence of the invariant propredicative clitics found to refer to predicative DP in (mainly canonical) copular sentences.

5. Propredicative clitics

Propredicative clitics are proforms found in copular sentences to refer to either the predicative NP (45) or to the AP (46). The propredicative clitics are invariant in gender and number, also when they refer to a feminine predicative NP as *la causa* 'the cause' in (45b) or to an inflected AP as *simpatiche* 'funny' (feminine plural) in (46b).

(45) a. Le ragazze sono la causa del litigio.
the girls be.3PL the cause of-the fight.
b. Le ragazze *la / lo sono [la causa del litigio][12]
the girls CL.3SG.F CL.3SG.N be.3PL the cause of-the fight
'The girls are the cause of the quarrel'

(46) a. Le ragazze sono simpatiche
the girls be.3PL funny
b. Le ragazze *le/ lo sono (simpatiche).
the girls CL.3PL.F CL.3SG.N be.3PL funny
'The girls are funny'

Although the accusative masculine object clitic is syncretic with the *lo* propredicative clitic, the former is inflected for gender and number (47), the latter is invariant (46).

(47) a. La ragazza la /*lo riconobbe [la zia]
the girl CL.3SG.F CL.3SG.M recognize.3SG the aunt
'The girl recognized her [the aunt]'

[12] The square brackets indicate that the phrase is not overtly realized: the clitic refers to the DP within the brackets.

b. La ragazza *la / lo riconobbe [lo zio]
The girl CL.3SG.F CL.3SG.M recognize.3SG the uncle
'The girl recognized him [the uncle]'

As in Italian, also in the Este variety the propredicative object clitic is realized as *eo*, which is invariant both in gender and number (masculine third person singular). For instance, in (48) the propredicative clitic *eo* refers to a singular feminine DP *ea causa dea barufa* 'the cause of the fight', in (49) it refers to an inflected (plural masculine) adjective *bei* 'beautiful'. The invariant propredicative clitic is syncretic with the accusative masculine clitic (50).

(48) a. I tosi ze ea causa dea barufa
the boys be.3PL the cause of-the fight
'The boys are the cause of the fight'
b. I tosi *ea$_i$ / eo$_i$ ze [ea causa]
the boys CL.3SG.F CL.3SG.N be.3PL the cause
Lit. 'The boys are it [the cause]'

(49) a. I tosi ze bei
the boys be.3PL beautiful.3PL.M
'The boys are beautiful'
b. I tosi *i$_i$ / eo$_i$ ze [bei]
the boys CL.3PL.M CL.3SG.N be.3PL beautiful
Lit. 'The boys are it [beautiful]'

(50) E tose eo ga visto
The girls CL3SG.M have.3SG seen
'The girls have seen him'

The same invariant propredicative clitic *eo* appears also in inverse copular sentences as in (51a-b). However, while the subject proclitics are available when a full predicate is in preverbal position (see 40 above), when the predicative NP is cliticized, it can never co-occur in a position higher than the subject clitics (51c-d) (as it was the case for *a* clitic, see 14-16) or can occur between a subject clitic and an inflected verb (51e). We argue that the two clitics share the same position within the inflectional layer. Although *eo* is invariant, it does not share any property with the higher *a* clitic (in the CP layer), since it is syncretic with masculine object clitic *eo*, which is within the inflectional layer.[13]

(51) a. Eo ze e tose [ea causa]
CL.3SG.N be.3PL the girls the cause
b. Eo ze i tosi [bei]
CL.3SG.N be.3PL the boys beautiful

[13] Manzini and Savoia (2005) account for the distribution of clitics proposing a string in which Definite clitics (subject clitics) are higher than 1st and 2nd person clitics (P) and 3rd person object clitics (N) which lexicalize N properties.

(i) [D…[P .. [N]]]

c.	*Eo	e	ze	e	tose	[ea	causa]
	CL.3SG.N	CL.3PL.F	be.3PL	the	girls	the	cause
d.	*Eo	i	ze	i	tosi	[bei]	
	CL.3SG.N	CL.3PL.M	be.3PL	the	boys	beautiful	
e.	*I	eo	ze		[i	tosi	]
	CL.3SG.N	CL.3SG.N	be.3PL		the	boys	

If we refer to the clitics within the inflectional layer (so excluding the *a* clitic in C) of nominal copular sentences, we find or subject clitics or invariant propredicative clitics. While the invariant *eo* clitic is found in both canonical and inverse sentences (see examples above in 48b, 49b, 50) when no other clitic is present and when the predicative NP is not overtly realized, inflected subject clitics are found in canonical sentences (see 40). However, when subject clitics appear in inverse copular sentences, they can only refer to the element targeted for agreement by the verb, i.e. the postverbal subject:

(52)	Ea causa	dea	baruffa$_j$	*ea$_j$/	*eo/	e$_i$
	the cause	of-the	fight	CL.3SG.F	CL.3SG.N	CL.3PL.F
	ze	e	tose$_i$			
	be.3PL	the	girls			

The status of subject clitic, in fact, is crucially linked to the verb inflection and to the element that is probed for agreement. However, subject and object clitic can co-occur with transitive verbs:

(53)	I	(e)o	gà	visto
	CL.3PL.M	CL.3S.M	have.3PL	seen
	'They have seen him'			

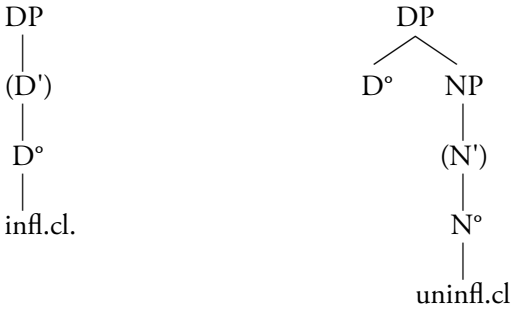
One question remained to be solved is why subject and propredicative clitic cannot co-occur. We argue that the answer is linked to the reduced structure of copular sentences, as proposed by Moro (1997), which is shared by other verbs which select a small clause as a complement. The verb *be* is merely inflection and selects a predicational phrase (small clause) as a complement in which the predication happens between a DP subject and a DP predicate. The clitic position available in a language with subject clitics such as the variety of Este (with the differences across the inflectional paradigm for number and person we have illustrated above) is used to lexicalize (optionally for 3rd person clitics) the D properties: that is, subject inflected clitics (not of the *a* type) are merely inflectional endings.

However, as for the invariant propredicative *eo*, we argue that predicate NPs can cliticize but they cannot show overt inflectional morphology. Predicative NPs (as sketched in section 4) introduce something like a description and have no referential status. Applying Moro's (1993, 1997) analysis for the invariant propredicative *lo* in Italian, we argue that the invariant *eo* in the Este variety should be analyzed as being generated in N0 rather than D0[14]:

[14] Recall that Moro (1997) shows (following the insight of Longobardi 1994) that the predicate NP can also be found with no determiners in nominal copular:

(i)	Le	ragazze	sono	(la)	causa	del	litigio.
	the	girls	be.3PL	(the.SG.F)	cause.SG.F	of-the	quarrel
	'The girls are (the) cause of the quarrel'						

(54) Moro (1997: 72)



Furthermore, the *eo* invariant clitic can be used only in strong pragmatic contexts, that is when the predicate NP can be inferred by the pragmatic context. Apparently this is a contrast with the subject clitic that represents mere inflection. However, we have seen that also subject clitics can be found with postverbal subjects (see 40 above) only in inverse copular sentences (with a topic-comment structure, where the predicate is the topic). We suggest that this general description may lead to the conclusion that clitic+copula represents always a configuration with comment interpretation (in a topic-comment configuration). While the invariant propredicative clitic refers to a description (non-referential) already presented within the discourse (recall that it cannot co-occur with overt predicative NP), the inflected subject clitic refers to the D properties of the focal postverbal subject. With this in mind, we try to sketch some syntactic considerations on the difference between canonical and inverse copular sentences that influence either the characteristics of agreement and the informational structure.

6. Syntactic analysis

All the considerations sketched above do not contradict in any respect the unified theory of copular sentences as proposed by Moro (1997). The syntax of nominal copular sentences interacts with the parametric difference linked on how the (D)efinite properties are lexicalized (that is, how we presented the *pro-drop* parameter in section 3). The N (predicative NP) versus D (subject NP) status of the NP may account also for the agreement differences in partial *pro-drop* languages where D properties can be lexicalized through subject clitics. The same syntactic configuration may imply different scope discourse semantic interface representations: namely the marked status of inverse copular sentences where a predicate NP introduces a description (N) and the postverbal NP specifies the entity that actually fits the description.

In sum, the background syntactic analysis coincides with Moro (1997): copular sentences involve a raising predicate that selects a small clause as its complement and does not have any particular meaning. The small clause is the place where the predication occurs between two XPs: a subject and a predicate. The subject generated in the small clause raises to the higher position of the copula as in (55), while the predicate stays *in situ*: this configuration was defined by Moro (1997) as the "canonical copular sentence".

(55) [[Una foto del muro]$_i$ [IP è [SC t$_i$ [la causa della rivolta]]]]
[[A picture of the wall]$_I$ [IP is [SC t$_i$ [the cause of the riot]]]]

The predicate generated into the small clause can raise to the higher position of the copula, while the subject DP stays *in situ* (56): this configuration is what Moro (1997) calls "inverse copular sentence".

(56) [[La causa della rivolta]$_k$ [IP è [SC [una foto del muro]t$_k$]]]
[[The cause of the riot]$_k$ [IP is [SC [a picture of the wall]t$_k$]]]

Among many other diagnostics, Moro (1997) noticed that languages may differ on the NP target of the agreement with the verb in inverse copular sentences. While in Italian the verb always agrees with the subject NP, in English the verb agrees always with the preverbal NP. Moro accounts for this difference through the presence of *pro* in the language. In inverse copular constructions, in Italian the predicate NP is in a position higher than IP, and the postverbal subject agrees since the *pro* is coindexed with it. In English, since no *pro* is available, the verb agrees with the only raised NP, i.e. the predicate NP.

In a partial *pro-drop* language, as the variety of Este, agreement works like in a *pro-drop* language (as Italian): the presence of *pro* determines agreement with the postverbal subject NP in inverse constructions. However, we have a specific agreement configuration for inverse copular sentences which is not found with other verbs: a subject proclitic co-occurs with a postverbal full DP (see example 40). Since subject clitics represent inflectional features, the use of subject clitic in inverse copular sentences needs to be linked to the pattern of agreement in inverse constructions.

If we reformulate the *pro-drop* parameter through differences across languages on how D properties are lexicalized, we argued that in the variety of Este the D properties are lexicalized either by the subject clitic or the inflectional morphology on the verb. However, this is not enough to account for the presence of a proclitic in inverse copular constructions. We have been pursuing that depending on person (1st person for saliency), quasi-argumental status of the NP bearing D properties (with meteorological verbs) and on the informational status of the the NP subject (right dislocated) there are differences in the overt realization of D properties. Therefore, comparing (57), where subject clitic is allowed, with (58), where subject clitic is forbidden, since we do not find any contrast linked to person or argumental status, we propose that in this case we are dealing with a contrast in the informational structure. While in (57), the predicative preverbal NP that defines the property of a subset that includes the postverbal NP appears before the verb, in (58) the predication is identified directly on the verb that bears agreement morphology.

(57) Ea causa dea barufa i$_k$ ze i so fradei$_k$
the cause of-the fight CL.3PL.M be.3PL the his brothers
'The cause of the fight is his brothers'

(58) (*Ea) ze rivà to mama.
CL.3SG.F be.3SG arrived your mum
'Your mum has arrived.'

In (57) there is an instantiation of a topic and the clitic + verb + NP is a *comment*, whereas the sentence in (58) can be interpreted as a broad focus sentence (used to answer to questions like 'What happened?'). The inverse copular sentences are informationally marked structures in which the preverbal predicative NP does not refer to an entity, but to a property representing the subset that includes the subject NP.

Following Moro (1997), who implements the account of Longobardi (1994), DPs are arguments and NPs are predicate, so the predicative NP defines a property and not an argument. Importantly, as mentioned in section 5 ff.14, NPs can be found without D determiner. The different status of the NP in inverse copular construction creates an informational marked sentence. Along this line, also the obligatory agreement with the subject in inverse copular constructions can be interpreted as an effect of the D versus the N status of the two NPs.

Following the minimalist analysis for AGREE (Chomsky 2001), the copula probes the nominal element that bears D features (as proposed for Persian and Armenian by Bejar and Kanehmuyipour 2017) and agrees always with the subject NP with D properties: the probe has specified property as its goal, in this case the D properties. Predicative NPs never have D features and cannot be targeted by agreement. In the variety of Este, the lower subject clitics (excluding the invariable *a*) are purely inflectional, so when the verb targets a NP with D properties, such properties are lexicalized by verbal morphology and/or subject clitics (depending on person, argumental status or informational structure).

As for the syntactic structure of both canonical and inverse copular sentences, what we proposed so far is perfectly compatible with Moro's analysis, the only difference is linked to how we analyze agreement if we use a probe/goal mechanism or a spec-head relation in INFL between a DP and a *pro*.

However, for the preverbal DP in copular sentences, we propose that both in inverse and canonical copular sentences the preverbal DP is higher than the specifier of INFL in the variety of Este. We follow the idea of Cinque (1979), Benincà and Cinque (1993) (among others) for which all subjects in Italian are dislocated. Since EPP and D are lexicalized by verbal morphology (or by subject clitics in partial *pro-drop* languages), the overt subject is realized when it has a clear scope discourse semantic effect ("criterial" in terms of Rizzi 2010). The high *a* clitic, which is in the C layer, is found after the preverbal NP and before the inflected subject clitic:

(59)	Ea	causa	dea	barufa	(a)	te	si	ti
	the	cause	of-the	fight	CL	CL.2SG	be.2SG	you.2SG
	'The cause of the fight is you'							

One more issue about the syntax of copular sentence which is confirmed by the data of the variety of Este is the reduced inflectional structure, as proposed by Moro (1997) to account for different phenomena including the presence of an invariable clitic to refer to the predicative NP. We have seen, in fact, that in the variety of Este (a part for the C clitic *a*) the inflected subject clitic and the invariant propredicative clitic compete for the same proclitic position (although subject clitics have D properties, while the propredicative clitics N properties). This is a further evidence of the fact that: (i) copular sentences (and probably also unaccusative verbs) have a reduced inflectional structure; (ii) the alternation between subject clitics and propredicative clitics depends mainly on the referential properties of the subject versus the predicative NP, and it is strictly correlated to the informational structure: in the variety of Este, propredicative clitics can only be found when the predicative NP is omitted.

7. Concluding remarks

In the present work we described the distribution of inflecting subject clitics, propredicative clitics, and higher invariant subject clitics in the variety of Este. Our description was focused on the distribution of clitics in copular sentences. Since nominal copular has always been rep-

resented as a challenge to the theory of predication because the predicate is nominal, we have described how clitic proforms interact with the syntax of copular constructions. In the variety of Este while subject clitics are allowed in canonical copular sentences, following the characteristics of the inflectional paradigm, with inverse copular sentences an overt subject clitic appears also with postverbal subject (configuration which is not allowed with lexical verbs). We analyzed this data as linked to the peculiar informational structure of the inverse copular sentences where the preverbal predicate NP represents a topic and the cluster subject clitic+copula+postverbal DP represents a comment to that topic. Since dislocation plays a crucial role in determining the appearance of subject clitics, we argued that the inflectional features encoded in subject clitics interacts with the informational structure. The co-occurence with the invariable *a* clitic confirms that while inflected subject clitics deal with the inflectional layer, *a* clitics deal with the C layer. The impossibility of finding inflected subject clitics in co-occurrence with invariant propredicative clitics confirms the reduced inflectional layer of copular sentences.

All the concluding remarks presented above do not contradict the unitary analysis of copular sentences presented by Moro (1993, 1997). The same syntactic structure can provide different interpretation at the semantic interface. In our perspective, the syntax of canonical and inverse sentences is mapped to informational interpretation that interacts with the overt realization of subject clitics.

Acknowledgements

We would like to thank Maria Rita Manzini for her preliminary feedback and two anonymous reviewers for the precious insights and suggestions. All mistakes are our own. A special thank goes to our native informants. This work is the result of the collaboration of the authors in all respects. Nevertheless, for Italian administrative purposes Linda Badan takes responsibility for Section 1, 2 and 5 as well as for the joint redaction of Section 7. Paolo Lorusso takes responsibility for Sections 3, 4, 6 as well as for the joint redaction of Section 7.

References

Béjar, Susana, and Kahnemuyipour Arsalan. 2017. "Non-canonical Agreement in Copular Sentences." *Journal of Linguistic*s 53 (3): 463-499. doi: 10.1017/S002222671700010X.

Benincà, Paola. 1994. *La variazione sintattica: Studi di dialettologia romanza*. Bologna: Il Mulino.

Benincà, Paola, and Guglielmo Cinque. 1993. "Su alcune differenze fra enclisi e proclisi." In *Omaggio a Gianfranco Folena*, ed. by Paola Benincà, and Guglielmo Cinque, 2313-2326. Padova: Editoriale Programma.

Bianchi, Valentina, and Adriana Belletti. 2016. "Definiteness Effect and Unaccusative Subjects: An Overview and Some New Thoughts." In *Definiteness effects: Bilingual, Typological and Diachronic Variation*, ed. by Susann Fischer, Tanja Kupisch, Esther Rinke, 14-65. Cambridge: Cambridge Scholars Publishing.

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

Bobaljik, Jonatan D. 2008. "Missing Persons: A Case Study in Morphological Universals." *The Linguistic Review*, special issue, *Examples of Linguistic Universals* 25 (1-2): 203-230. doi: 10.1515/TLIR.2008.005.

Brandi, Luciana, and Patrizia Cordin. 1981. "Dialetti e italiano: un confronto sul parametro del soggetto nullo." *Rivista di grammatica generativa* 6: 33-87.

Calabrese, Andrea. 2011. "Investigations on Markedeness, Syncretism and Zero Exponence in Morphology." *Morphology* 21 (2): 283-325. doi: 10.1007/s11525-010-9169-y.

Cardinaletti, Anna, and Lori Repetti. 2008. "The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects." *Linguistic Inquiry* 39 (4): 523-563.

Cardinaletti, Anna, and Lori Repetti. 2010. "Functional Vowels in Main Questions in Northern Italian Dialects." In *Romance Languages and Linguistics Theory 2008: Selected papers from 'Going Romance' Groningen 2008*, ed. by Reinecke Bok-Bennema, Brigitte Kampers-Manhe, and Bart Hollebrandse, 37-58. Amsterdam: John Benjamins.

Chomsky, Noam. 1981. *Lectures on Government and Binding. The Pisa Lectures*. Dordrecht-Cinnaminson, NJ: Foris.

Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: A Life in Language,* ed. by Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/4056.003.0004.

Cinque, Guglielmo. 1979. "Left Dislocation in Italian: A Syntactic and Pragmatic Analysis." In *Studi di sintassi e pragmatica*, ed. by Guglielmo Cinque, 97-146. Padova: Clesp.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non-)veridical Dependency.* Amsterdam: John Benjamins.

Harley, Heidi, and Ellen Ritter. 2002. "Person and Number in Pronouns: A Feature Geometric Analysis." Manuscript, University of Arizona and University of Calgary.

Holmberg, Anders. 2010. "Null Subject Parameters." In *Parametric Variation: Null subjects in Minimalist Theory*, ed. by Biberauer, Theresa, Anders Holmberg, Ian Roberts and Michelle Sheehan, 88-124. Cambridge: Cambridge UP.

Holmberg, Anders, and Michelle Sheehan. 2010. "Control into Finite Clauses in Partial Null-Subject Languages." In *Parametric Variation: Null subjects in Minimalist Theory*, ed. by Biberauer, Theresa, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 125-152. Cambridge: Cambridge UP.

Huang, C.-T. James. 1984. "On the Distribution and Reference of Empty Pronouns." *Linguistic Inquiry* 15 (4): 531-574.

Longobardi, Giuseppe. 1994. "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form." *Linguistic Inquiry* 25 (4): 609-665.

Manzini, M. Rita. 2012. "From Romance Clitics to Case: Split Accusativity and the Person Case Constraint." In *Romance Languages and Linguistic Theory 2010: Selected Papers from 'Going Romance' Leiden 2010*, ed. by Irene Franco, Sara Lusini, and Andrés Saab, 1-20. Amsterdam: John Benjamins.

Manzini, M. Rita. 2015. "On the Substantive Primitives of Morphosyntax and their Parametrization: Northern Italian Subject Clitics." In *Representing Structure in Phonology and Syntax*, ed. by Marc van Oostendorp, and Henk van Riemsdijk, 167-194. Berlin: De Gruyter Mouton. doi: 10.1515/9781501502224-007.

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

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

Manzini M. Rita, and Leonardo M. Savoia. 2011. *Grammatical Categories: Variation in Romance Languages*. Cambridge: Cambridge UP.

Manzini M. Rita, and Kenneth Wexler. 1987. "Parameters, Binding Theory, and Learnability." *Linguistic Inquiry* 18 (3): 413-444.

Moro, Andrea. 1993. *I predicati nominali e la struttura della frase*. Padova: Unipress.

Moro, Andrea. 1997. *The Raising of Predicates*. Cambridge: Cambridge UP.

Nicolis, Marco. 2008. "The Null Subject Parameter and Correlating Properties: the Case of Creole Languages." In *The Limits of Syntactic Variation*, ed. by Theresa Biberauer, 271-294. Amsterdam: John Benjamins.

Poletto, Cecilia. 1993. *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. Oxford-New York, NY: Oxford UP.

Raposo, Eduardo, and Juan Uriagereka. 1996. "Indefinite *se*." *Natural Language and Linguistic Theory* 14 (4): 749-810.

Renzi, Lorenzo, and Laura Vanelli. 1983. "I pronomi soggetto in alcune varietà romanze." In *Scritti linguistici in onore di Giovan Battista Pellegrini*, vol. I, ed. by Paola Benincà, 391-419. Pisa: Pacini.

Rizzi, Luigi. 1986. "On the Status of Subject Clitics in Romance." In *Studies in Romance Linguistics*, ed. by Osvaldo Jaeggli, and Carmen Silva-Corvalan. 391-419. Dordrecht: Foris. doi: 10.1515/9783110878516-025.

Rizzi, Luigi. 2010. "On Some Properties of Criterial Freezing." In *The Complementizer Phase: Subjects and Operators*, vol. 1, ed. by E. Phoevos Panagiotidis, 17-32. Oxford: Oxford UP.

Roberts, Ian. 2010. *Agreement and Head movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.

Roberts, Ian. 2014. "Subject Clitics and Macroparameters." In *Diachrony and Dialects. Grammatical Change in the Dialects of Italy*, ed. by Paola Benincà, Adam Ledgeway, and Nigel Vincent, 177-201. Oxford: Oxford UP.

Sportiche, Dominique. 1992. "Clitic Construction." Manuscript, UCLA.

Sportiche, Dominique. 1996. "Clitic Constructions." In *Phrase Structure and the Lexicon,* ed. by Johan Rooeyck and Louise Zaring, 213 -277. Dordrecht: Kluwer.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: B. Baldi, L.M. Savoia (2021) Object clitics for subject clitics and DOM phenomena in the Franco-Provençal dialects of Apulia. *Qulso* 7: pp. 33-49. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12001

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Object clitics for subject clitics and DOM phenomena in the Franco-Provençal dialects of Apulia*

Benedetta Baldi and Leonardo M. Savoia
Università degli Studi di Firenze
(<benedetta.baldi@unifi.it>; <lsavoia@unifi.it>)

Abstract:

The relation between auxiliaries and argumental properties of the lexical verbs is the question raised by the facts we will investigate. Other issues are posed by the role and the behavior of deictic clitic pronouns in different contexts and in the clitic string. The data examined in this contribution come mainly from the Franco-Provençal dialects spoken in the villages of Celle San Vito and Faeto in North Apulia; these dialects are the heritage linguistic use of ancient minorities dating back to the thirteenth century. We will briefly compare these data with the related data of the Franco-Provençal varieties of West Piedmont. The analysis we propose relies on the idea that auxiliary verbs are not functional elements but preserve their lexical properties, as shown by the identic syntactic behavior of *be* in its different occurrences. The theoretical frame we follow is the formulation proposed in Chomsky (2020a, b), based on the merge operation and the Labeling Algorithm, leading to a more appropriate treatment of "head movement" and agreeing syntactic structures.

Keywords: *Auxiliary, Copula, DOM, OCls for SCls, Participle*

1. The distribution of SCls and OCls in Apulian Franco-Provençal

In the villages of Celle and Faeto (North Apulia) a Franco-Provençal dialect is spoken, brought here by the people that, approximately in the 13th century, founded these settlements.

*The data analyzed in this article were collected through field investigations with native speakers. We especially thank Agnesina Minutillo of Celle and Cocco Maria Antonietta and Giovanni Marella of Faeto. The data from Cantoira have been provided to us by our informants Bruna and Martina Ravicchio.

The contribution of the authors can be summarized as follows: Conceptualization, B. Baldi and L.M. Savoia; methodology, B. Baldi; theoretical framework, L.M. Savoia and B. Baldi; data curation, L.M. Savoia; writing - original draft preparation, L.M. Savoia and B. Baldi, review and editing, B. Baldi.

Apulian Franco-Provençal shows a system of pronominal clitics (Subject Clitics and Object Clitics) well documented for the dialects of their original area (Southern France, Aosta Valley and West Piedmont). Leaving aside for the moment the latter, we focus on the Apulian Franco-Provençal, where we find the paradigm for SCls in (1a) and for OCls in (1b), generally distinct except for the 2nd person, where the single common form *tə*/*və* 'you' is attested. As can be seen in (1a, b), the dative form is missing, and *sə* is the reflexive and includes the 1PL as well. In (1c) the stressed strong forms of personal pronouns are provided, for subject and object contexts. Only in the case of the 1st singular person an oblique form is present. In other cases, the same form occurs as the subject and the object of a verb or a preposition. As generally in Northern Italian systems, SCls can combine with expressed subjects, including the stressed pronouns.

(1)

a.

	SCl
1st	dʒə
2nd	tə
3rd	i
1st pl	nə/ i
2nd pl	və/ i
3rd pl	i

b.

	OCl	Refl
	mə	(mə)
	tə	(tə)
	l-u / l-a	sə
	sə	sə
	və	sə
	l-o / l-ə	sə

c.

			Oblique	
1st	ddʒi	'I'	mi	'me'
2nd	ti	'you'		
3rd	ijə / iʎʎə	'he/ she'		
1st pl	nus	'we'		
2nd pl	vu	'you'		
3rd pl	isə	'they'		

Celle

With this scheme in mind we are able to illustrate the distribution of SCls and OCls in sentences. SCls are exemplified for unergatives in (2a) and for unaccusatives in (2b). Transitives with postverbal objects have the same distribution, as in (2c).

(2)

a

ddʒə ddɔ:rə
tə ddɔ:rə
i ddɔ:rə
nə durmuŋŋə
və durmijə
i durmundə
'I slip, you slip,…'

b.

ddʒə viŋŋə
tə viŋŋə
i vində
nə vənuŋŋə
və vənijə
i vənundə
'I come, you come, …'

c.

ddʒə vajə (tuttə kwandə)
tə vajə
i vajə
nə viuŋŋə
və vijijə
i viundə
'I see everybody, you see everybody, …'

As in many Northern Italian dialects (Manzini and Savoia 2005, 2011), OCls exclude SCls except 1st and 2nd singular persons, *dʒə* 'I' and *tə* 'you' respectively, as in (3i) and (3ii). In the case of a 2nd person subject and a 1st person object the reversed order appears and the SCl is realized by the objective form *mə*, as in (3ii.a), (3iii-vi) illustrate the exclusion of SCls by OCls in the other persons. The coinciding forms *tə*/*və* 'you' are specified as '2SG/PL'. Subjects are labeled S and objects O, and the six persons are indicated by Roman numerals.

(3) i. ddʒə l-u / tə vajə
S.1SG O.3-MSG / 2SG I.see
'I see him/ you'

ii. ttə l-u / l-a / sə vajə
2SG O.3-MSG / O.3-FSG / 1PL you.see
'you see him/ her/ us'

ii.a mə tə vajə
O.1SG 2SG you.see
'you see me'

iii. l-u / mə / tə vajə
O.3-MSG / O.1SG / 2SG (s)he.sees
'(s)he sees him/ me/ you'

iv. l-u / tə viuŋŋə
O.3-MSG / 2SG we.see
'we see him/ you'

v. l-u / mə vijijə
3-MSG / 1SG you.see
'you(pl) see him/ me'

vi. mə / tə / l-a viundə
O.1SG / 2SG / O.3-FSG they.see
'they see me/ you/ her'

Celle

In the verbs with three arguments like *to give*, the proclitic order is the same as that of Italian, where the 1st/2nd OCls precede the 3rd person OPCls, as in (4a, a'). Again, SCls are excluded, except 1st/2nd singular elements, and, the 1st person OCl precedes the SCl in the combination *2nd person SCl+1st person OCl*, as in (4b); see also (3ii.a) above:

(4) a. (ijə) mə l-u dde:nə
he me O.3-MSG he.gives
'He gives it to me'

a'. (ti) tə sə l-u dde:nə
you you us O.3-MSG you.give
'You gives it to us'

b. (ti) mə tə l-u dde:nə
you me you O.3-MSG you.give
'You give it to me'

Celle

Reflexive/non-active forms exclude the specialized SCLs. We find the objective forms *mə* and *tə* for the 1st/2nd singular persons in (5i, ii) and the reflexive *sə* in all other forms. in (5iii-vi). So, 1st and 2nd person SCls are in turn prevented from combining with OCls.

(5) i. mə lla:və
ii. tə lla:və
iii. sə lla:və
iv. sə lavuˈə
v. sə la'va
vi. sə lavundə
'I wash myself, you wash yourself, …'

Celle

The compatibility between clitics, as we noticed, is constrained by person, so that the only strings allowed are those in (6), where SCl precedes an OCl, limited to 1st/2nd person SCls.

(6)

SCl	+	OCl
1st	dʒə	OCLs
2nd	tə	OCls

Other combinations SCl+OCl(s) are not instantiated, considering that reflexive contexts in turn exclude all SCls.

1.1 DOM effects

Let us dwell now on the OCls of 1st and 2nd person, observing their occurrence in the sentence. In Apulian Franco-Provençal the paradigm of OCls does not include specialized forms for dative (vs. accusative). In the case of 1st and 2nd person forms this syncretism is present in most Romance varieties, whereby the same form occurs as the direct object, in (7a) and dative in (7a') (cf. Manzini *et al.* 2020). As for the 3rd person, we find that the accusative form occurring in transitive contexts in (7b, b') is used in dative contexts, with the consequence that the *3rd dative+accusative* reading is realized only by the accusative form corresponding to the IA/Object, as in (7c), a type of distribution attested in some Southern Italian dialects (Manzini and Savoia 2005).

(7) a. mə / tə vi-undə
me / you see-3PL
'they see me/ you'
sə / və viundə
us/ you see-3PL
'they see us/ you'

a'. mə / tə / sə / və denə sett-a tʃuɔsə
me / you / us/ yo give-3SG this-FSG thing
'they give to me/ you/ us/ you this thing'

b. l-u/ l-a/ l-ɔ/ l-ə vi-undə
3-MSG/-FSG/-MPL/-FPL see-3PL
'they see him/ her/ them'

b'. mə / tə / sə / və l -u/ l-a/ l-ɔ/ l-ə dən-undə
me / you / us/ you 3-MSG/-FSG/-MPL/-FPL give-3PL
'they give it/ them to me/ you/ us/ you'

c. ijə l-u/ l-a/ l-ɔ/ l-ə denə
he 3-MSG/-FSG/-MPL/-FPL give-3SG
'he gives it (MSG/FSG)/ them to him/her/them'

Celle

This distribution suggests that, while 3rd person clitics are accusative and realize the IA, 1st and 2nd forms are oblique, in so far as they realize the dative reading, as in (7a'). In other words, a DOM system shows up which contrasts the discourse participants with the participants to the event expressed by the verb. In the case of stressed personal pronouns, we note that in object contexts they are always governed by a preposition, including the object of a transitive, confirming the DOM treatment, as in (8a). The same forms occur in combination with the other prepositions, as in (8b). All the pronouns in (8a,b) can occur as subjects, except for the first person, which distinguishes *mi* from the subject element *ʤi*, in (8c).

(8) a. i camm-undə a mmi /tti/ ijə / iʎʎə / nnus / vvu / isə
SCl.3PL call-3PL to me/ you/ him/ her/ us / you/ them
'they call me/you/ him/ her/ us / you/ them'

b. pə / dɔ mmi /tti/ ijə / iʎʎə / nnus / vvu / isə
for / with me/ you/ him/ her/ us / you/ them

c. ʤ-i ʤə dɔr
1-INFL SCl sleep.1SG
'I sleep'

Celle

It is natural to see in the prepositional accusative the result of the long-lasting contact between Franco-Provençal and Apulian surrounding dialects. Contact has brought about important re-organizations in Franco-Provençal systems, as discussed in Baldi and Savoia (2021). As for prepositional objects, definite IAs are introduced by the preposition *a*, on a par with the other Southern Italian dialects.

According to Manzini *et al.* (2020), DOM is a manner of embedding sub-sets of definite arguments of which the core-set is represented by personal pronouns and, within them, 1st and 2nd persons. In the typological literature the behaviour of 1st and 2nd person in comparison with 3rd person pronouns/demonstratives and NPs is treated in terms of referential properties (animacy or definiteness) expressed by means of a hierarchy regulating the distribution of grammatical functions in case systems (Dixon 1994, Kiparsky 2008).[1] Kiparsky (2008: 34) associates the referential hierarchy with definiteness, as the property that anchors the syntactic role of the nominal and pronominal elements. Languages vary by cutting the definiteness hierarchy at different levels, including all definite NPs, only kinship terms in possessive constructs and personal pronouns, only personal pronouns or, finally, only 1st/2nd person elements. The crucial insight is that a Participant internal argument reflects a complex organization of the event "where the Participant plays the role of possessor (locator) of the VP event" (Manzini *et al.* 2020: 242). Our idea is that "the surfacing of highly ranked (DOM) internal arguments as datives is no morphological accident […] DOM arises in the syntax, reflecting a slightly different structuring of the event with Participant internal arguments". However, differently from prepositional datives with *give*-type verbs, where the preposition is required by the predicate, in the case of DOM the preposition is required by the referential properties of the IA, which is treated as a sort of possessor of the event. In other words, in this perspective, DOM is not simply a surface fact, but corresponds to a structural obliquization of IA when being a high ranked referent.

[1] The referential hierarchy in (i) (Dixon 1994: 85; Kiparsky 2008: 34)

(i) 1P > 2P > 3P/ demonstratives > proper nouns/ kin terms > human > animate > inanimate

accounts for the relation between nominal elements and grammatical functions, whereby "a first person pronoun is more likely than any other NP constituent to be in A [subject of transitives] rather than in O [object of transitives] function. Next most likely as A is second person pronoun, then demonstratives and third person pronouns, followed by proper names" (Dixon 1994: 85).

If we follow Fillmore (1968) in assuming that cases are the inflectional equivalent of prepositions, the elementary introducers such as Italian *di/a* or English *of/to* would be equivalent to the genitive and the dative respectively. The oblique case, on a par with a preposition, is a predicate introducing a relation between the argument it selects and another argument. This means that we assign a relational content to cases, with the effect that the oblique case or prepositions are endowed with interpretable properties. This solution contrasts with the more traditional view that prepositions like 'of' or 'to' are devoid of interpretive content, or, in minimalist terms, uninterpretable (Manzini *et al.* 2020).

The idea we adopt, going back to Manzini and Savoia (2011), is that oblique case and prepositions such as 'of' or 'to' can be thought as elementary part/whole relators; in other words, a single property, namely inclusion/superset-of, formalized as [⊆], is associated with the conceptual cluster underlying oblique and/or of/to-like prepositions. This proposal is based on an insight present in the literature, whereby possession is the manifestation of the part-whole relation (Belvin and den Dikken 1997: 170). So, in the structure in (9), a preposition like Italian and Romance *a* (or English *to*), endowed with (⊆) content, takes as its internal argument its sister NP *ijə* 'he' and as its external argument the sister to its projection, i.e. the NP *lu livrə* 'the book'. The second internal argument of 'give', i.e. the dative, contributes to fixing the reference of the first internal argument, i.e. the accusative, by denoting a superset/domain/zone including it.

(9) [v də′na [PredP [DP lu livrə [PP⊆ [P⊆ a [DP ijə]]]]

Manzini and Savoia (2014), Manzini and Franco (2016) extend this analysis of dative/genitive as elementary relators, to DOM, so that the oblique of person in a transitive context will be characterized by the same structural representation, as in (10)

(10) [T i cammundə [VP … [PP⊆ [P⊆ a [DP mi]]]] 'they call (to) me'

A connected question is the form of the pronoun, that, as seen, in the case of 1st singular element differentiates the subject *ddʒ-i* form and the oblique *m-i*. Taking into account the preceding discussion, we can think that *m-i* is lexically specified by the oblique property [⊆], as in (11).

(11) [Infl [√ m [⊆]] -i]

In the case of all the other pronouns we find a single form for subject and oblique contexts, suggesting that this specialized specification has been lost. This recalls the paradigms of Northern Italian dialects where all the stressed pronouns have a single form for subject and object contexts.

2. Auxiliaries

Auxiliary selection separates unaccusatives and reflexives, with *be*, from unergatives and transitives, with *have*. While the latter have the same SCl system as the lexical verbs, in (12a,b), unaccusatives and reflexives select *mə* in SCl position, in (12c,d). As we can expect, passives match the other contexts with *be*, as in (12e).

(12) a.

dʒ e(nnə) dur'mi	
t a(nnə) dur'mi	
i attə dur'mi	
n a'vuŋŋə dur'mi	
v a'vi dur'mi	
i andə dur'mi	
'I have slept, …'	

b.

dʒ	e	la′va	(lo ddra)
t	annə	la′va	
i	attə	la′va	
n	avuŋŋə	la′va	
v	avijə	la′va	
i	antə	la′va	
'I have washed the clothes, …'			

c.	mə sejə la'va	d.	mə sejə və'ni
	tə sejə la'va		tə sejə və'ni
	s ettə la'va		ʌ ettə və'ni
	sə suŋŋə la'va		nə suŋŋə və'ni
	sə si la'va və		və si və'ni
	sə sundə la'va		i sundə və'ni
	'I have washed myself,…'		'I have come, …'

c'. mə sejə krəvɛrə / krəvɛrə-tə
me I.am covered.MSG/ covered-FSG
'I have covered myself'

e. mə sejə sta cam′ma da isə
i ettə sta cam′ma da isə…
'I have been called by him, she has been called by him, …'

Celle

Interestingly, *mə* occurs instead of the SCl *dʒə* in the case of predicative constructs with *be*, regardless the individual- or stage-level interpretation, and with *stay/stand*, as illustrated in (13a) and (13b) respectively. We note that the 3rd singular person of *be* has two alternants, *e* and *ettə*.

(13) a. mə sejə kun'teŋŋə / autə
tə sejə kun'teŋŋə
ʌ e(ttə) kun'teŋŋə
nə suŋŋə kun'teŋŋə
və si kun'teŋŋə
i sundə kun'teŋŋə
'I am glad/ tall, …'

b. mə/dʒə stə kə'tʃa / kundeŋŋə də ti
me/ S.1st stay laid out / happy of you'
'I am lying / happy with you'
mə/dʒə stə i'ki
me/S.1st stand here
'I am here'

Celle

stay as the progressive auxiliary selects, in turn, *mə*. This holds not only in unaccusative/reflexive contexts, in (14a), but also with agentive constructs, as in unergative and transitive contexts in (14b) and (14c).

(14) a. m əstə vənaŋŋə / lavaŋŋə
me stay.1SG coming / washing
'I am coming / washing myself'

b. m əstə durmaŋŋə
me stay.1SG sleeping
'I am sleeping'

c. mə l əstə faʃaŋŋə
me it stay.1SG making
'I am making it'

The insertion of *mə* for *dʒə* with *be* and *stay* in (13) and (14) contrasts with the behaviour of SCls in modal and deontic periphrases, where the occurrence of *mə* for *dʒə* depends on the obligatory 'clitic climbing' on the modal/ deontic verb. As a consequence, we find *mə* (for *dʒə*) in reflexive contexts, (14'a), and in modal contexts also with the auxiliary *be*, as in (14'b). As to the deontic periphrasis, *have* selects the infinitive introduced by the preposition *a* 'to', as in (14'd) so that we find *mə* only in the case of the reflexive reading. Otherwise, *dʒə* occurs, as in (14'c).

(14') a. mə vuəʎʎə la′va
me I.want wash
'I want to wash myself'

b. mə sejə pa pu′ti vənijə
me I.am Neg can.PPrt come
'I could not come'

c. dʒ e pa u′li durmijə
I I.have Neg wanted sleep
'I did'nt want to sleep'

d. m (e) a lla′va / dʒ (e) a ddurmijə
me (have) to wash / I (have) to sleep
'I have to wash-myself/ I have to sleep'

Celle

The variety spoken in Faeto is substantially identical to that of Celle, we could say the same one, except for some minor morphosyntactic and phonological discrepancies. As regards the phenomenon in question, we find the same distribution of SCls and OCls as in Celle, but the auxiliary is characterized by a partially different inflectional paradigm and, especially, by the single 1st person form ɛ for all verbal classes, as in (15a-d). *mə* for *dʒə* occurs also with stay as introducer of a gerund in (15e) and lexical verb in (15f).

(15) a. dʒ ɛ dur'mi
t a dur'mi
i attə dur'mi
a'vunnə dur'mi
v a'vi dur'mi
i andə dur'mi
'I have slept, …'

b. dʒ ɛ la′va (lo ddra)
t a la′va
i attə la′va
n avunnə la′va
v avi la′va
i antə la′va
'I have washed the clothes, …'

c. m ɛ la'va
t ɛ la'va
s ɛ(ttə) la'va
nə suŋŋə la'va
və si la'va və
sə sundə la'va
'I have washed myself,…'

d. m ɛ və'ni
t ɛ və'ni
ʎ ɛ və'ni
nə sunnə və'ni
və si və'ni
i sundə və'ni
'I have come, …'

e. mə l əstə lə′ʃannə
me it stay.1SG reading
'I am reading it'

f. m əstə ik′ki
me stand here
'I am here'

Faeto

Moreover, this form occurs also in the 1st person of *be* in predicative contexts, as in (16a, b), where, according to a general contrast registered in Southern Italian dialects, (16a) illustrates the individual-level predicates with *be*, while (16b) illustrates the stage-level predicates with *stay*. This suggests that *ɛ* does not belong to the paradigm of *have* but is an exponent of *be*.

(16)	a.		b.	
		m ɛ autə		m əstə kundennə
		t ɛ autə		t əstə kundennə
		ʌ ɛttə autə		ʌ əstə kundennə
		nə sunnə autə		nə stunnə kundendə
		və si autə		və stijə kundennə
		i sundə autə		i sundə kundennə
		'I am tall, etc.'		'I am glad, etc.'

Faeto

In summary, in these dialects, the 1st person displays the distribution in (17):

(17) 1st person SCl

transitive/ unergative	unaccusative	have	reflexive	be/ stay
dʒə	*dʒə*	*dʒə*	*mə*	*mə*

We conclude that auxiliaries have the same selectional properties as lexical verbs, as we may expect if auxiliaries are properly verbs, i.e. they correspond to an autonomous vP projection.

3. A comparison with the Piedmontese Franco-Provençal

The distribution of SCls sensitive to the context of occurrence is also attested in Franco-Provençal dialects of Western Piedmont (Roberts 1993, 2018, Manzini and Savoia 2005, 2010). In these dialects, the 1st person SCl *dʒə/i* 'I' does not combine with 3rd person OCls as in (18). (18a) and (18a') show the exclusion of the 1st person subject clitic *dʒ* in the presence of 3rd person OCls. This happens independently from the phonological context, i.e. we find the vocalic allomorph *dʒi* before OCls beginning both with a consonant and a vowel, as in (18b, b'). (18c, d) illustrate the lack of the clitic before *be*, auxiliary and copula respectively.

(18)	a.	l-u / lə / l-i / əl	tʃam-u
		OCl-MSG / OCl.FSG / OCl.MPL / OCl.FPL	call-1SG
		'I call him/ her/ them'	
	a'.	l-u	tʃa'meŋ
		him	call-1PL
		'We call him'	
	b.	dʒi t / u	tʃam-u
		SCl.1SG OCl.2SG / OCl.2pl	call-1SG
		'I call you.sg/pl'	
	b'.	dʒi t	tʃa'meŋ
		SCl.1sg OCl.2sg	call-1PL
		'We call you'	
		dʒi t eŋ	tʃa'ma
		SCl.1P OCl.2PSG have.1PL	called
		'We have called you'	

c.	se	/ seŋ	vy'ny	
	be.1SG	/ be.1PL	come	
	'I have / we have come'			
d.	ser-u	/ ser-unt	kuntənt	
	be.IMPF.1SG	/ be.IMPF.1PL	glad	

Cantoira

Overall, this distribution resembles that of Celle and Faeto, both in the exclusion when 3rd person OCls occur, and before *be*. Reflexives in Piedmontese Franco-Provençal select *have*, and therefore they are aligned with transitives. It turns out then that this particular distribution of the 1st person SCls stems from the original system of these varieties, being shared by non-adjacent dialects belonging to the same group, and is not due to the effect of changes caused by contact in the Apulian Franco-Provençal.

4. The analysis: auxiliaries and participles

The assumption that auxiliaries are the morphological exponents of functional heads is very familiar in the generative framework, starting from Chomsky (1957), Burzio (1986), a simple solution giving an answer to the problem concerning the fact that the auxiliary verb and the participle denote a single event (Manzini and Savoia 2011). Bentley and Eythórsson (2004: 447) treat "perfective auxiliaries as morpho-syntactic markers of tense and aspect". Their idea is that the insertion of *be* or *have* is triggered by certain sub-set of the semantic features associated to the verbal classes. As to the participle there are different proposals, as a reduced verbal structure (Belletti 1990, 2005) or a structure endowed with the entire set of the functional projections of V (Kayne 1993). The systems in which participles can take enclitic pronouns seem to argue for the latter conclusion, or, however to support the autonomous nature of participles.

More in general, *be* is also the verb of copular sentences and *have* can occur as a full verb of possession and in deontic phrases, and an adequate theory of *be* and *have* should treat their different occurrences in a unified way. D'Alessandro and Roberts (2010: 50 and ff.) assume that auxiliaries are raising verbs, substantially in line with Moro (1997) for *be*, and Manzini and Savoia (2005: 547). Manzini and Savoia (2011: 222, 223) characterize auxiliaries as full verbal projections embedding the sentential domain of a lexical verb:

> […] a relevant observation is that in all Romance and Albanian varieties, the *be* auxiliary of the perfect is also the copula […] This has a single argument slot, that of the embedded predicate […], which becomes associated with the matrix EPP argument (represented by the finite inflection of the copula). […] the auxiliary selection patterns according to transitivity and/or voice can be described by saying that the 'defective' *be* is restricted to selecting 'defective' – i.e. intransitive and/or middle-passive – predicates; conversely, the transitive ctive *have* is restricted to selecting transitive and/or active predicates.

Chierchia (2004: 47) connects the insertion of *essere* 'be' in languages such as Italian to the semantic properties of the lexical verb:

> Since auxs are property modifiers, they can be sensitive to the semantic make-up of what they modify. […] The choice of essere vs. avere in Italian, on the other hand, is sensitive to a different factor: subject affectedness. More explicitly, the range of subject-affecting operations […] constitutes the domain of essere;

Subject affecting operations include passives, reflexives and unaccusatives, all affecting the external argument by existentially quantifying over it (passives) or identifying it with the object (reflexive and unaccusatives). The result is that an argument, specifically the external one, has no morpho-phonological realization by the syntax, which, on the contrary, shows agreement properties identifying the IA as the subject of the clause. Along the same lines, in Manzini *et al.* (2016: 146) the contrast between *be* and *have* is related to a difference in their selectional properties, whereby *be* selects a "reduced" argument structure while *have* selects predicates with a closed argument structure, "in the sense that no free variables or generically closed ones are instantiated within it". The crucial role is played by the adjectival nature of past participle, selecting only the IA slot. Taking Manzini *et al.* (2015), and D'Alessandro and Roberts (2010) into account, our analysis is inspired by the idea that the past participle in Romance varieties and, generally, in Indo-European languages, is nothing but an aspectual item with an adjectival nature, giving rise to a clause including one argument.

4.1 A minimalistic approach to the auxiliary-participle *periphrasis*

In what follows we will analyze the auxiliary periphrases in the light of the recent proposals of Chomsky (2020a: 50, 51) based on the notion of (pair-)merge procedure and modification. Modification as in the case of an adnominal adjective expression such as *young man*, is the result of an operation of conjunction, that "contains elements, each of which is predicated of something. So we have a sequence of elements that looks like [19], with links L_i".

(19) $< \text{CONJ}, < S_1, L_1 >, \ldots, < S_{n,\ Ln > >}$

Along these lines, Chomsky (2020a: 51) identifies the Link with the categorizers *n* and *v*, with which R(oots) merge. We can conceptualize the categorizers *n* or *v* as the bundles of φ-features that characterize the functional content of words entering into the agreement operations. Drawing on Manzini *et al.* (2020) and their preceding work on nominal inflection, it seems natural to assume that *n* is the label for the class and number features of nominal agreement. In other words, in the case of items including nominal inflectional properties, as nouns and adjectives/participles, φ is the bundle of features corresponding to the nominal properties (see also Roberts 2018). Chomsky (2020a: 55) sees in pair-merge the way of treating head raising: "It's always described incorrectly. If a verb raises to inflection, say to T, it's always described as if the T-V complex becomes a T; but it's not, it's a V-the outcome of the adjunction is really verbal, not inflectional".

Consider now the auxiliary periphrases with *be* and *have*. We start from the elementary hypothesis that auxiliaries are full verbal projections, embedding a predicative relation between a noun and a participle selecting it as IA (Manzini and Savoia 2011). In the case of non-active *be* constructs, their coincidence with the copular constructs provides clear evidence in favour of the idea that the past participle and its argument form a small clause. More precisely, we treat the participle as a nominal form including a category-less lexical root R√ (in Manzini and Savoia 2017, Savoia *et al.* 2019), combining with the φ-features endowed with interpretive content (nominal gender, number properties); applying a proposal of Manzini and Savoia (2005) we associate the thematic vowel with a category of nominal class. In (20a,b) the analysis of Celle gender alternating past-participles such as *krəvɛrə* 'covered.MSG'/ *krəvɛrə-tə* 'covered.FSG' (cf.

(12c')), and of the invariant forms such as *la'v-a* 'washed' (cf. (12c)), is provided. We identify the element *-tə* in (20a) with Gender and the thematic vowel *-a* in (20b) with the exponent of Class.[2]

(20) a. [[krəvɛrə $_{R}$] tə $_{Gender}$]
b. [[lav $_{R}$] a $_{Class}$]

If the past participle has the properties of nouns and adjectives, as we assume, it selects only one argument. This can be due to its stative/resultative nature implying as a property-bearing referent, typically the IA. The lexical root can include agentivity, even if the EA is not realized by the structure hence causing some ambiguity, only solved by inserting an agentive complement.

In other words, *be*, both copula and auxiliary has the same lexical properties, embedding a predicative relation between a noun and an adjective/participle implementing the lexical properties of the verb, i.e. theta-roles associated with the root R. As a result, a sequence is yielded by pair-merging the clitic *mə* with the participle, on the basis of agreeing φ-features (Chomsky 2020a), as in (21). The referential properties of the1st and 2nd person, *mə*/*tə*, are compatible with the φ-features of the participle, to which their deictic specification is added.

(21) <mə, [[krəvɛrə $_{R}$] tə $_{Class/gender}$]>

Therefore, *be+adjective/participle* gives as a result a sentence in which the subject is the only argument of the nominal/participial item, i.e. the IA. *mə* and *tə* are merged with *v* and then with T, fixing the agreement properties of the verbal head, as in (22)[3]. This analysis closely recalls the proposal of Roberts (2010, 2018), whereby the Romance OCls can be understood as the morpho-phonological realization of the agreement φ-features of *v*.

(22) CP [<mə, T-sejə >[$_{vP}$ v [<R_x, PPrt/n> ($mə_x$)]]

Relying on the discussion in Chomsky (2020b: 166), we can conclude that [$_{C}$ mə- [T]] gives rise to a labelled construction "by shared and agreeing φ-features", where T inherits the features from C and *mə* is the only available nominal element (see also Roberts 2018)[4]. The same is true for copular contexts, where *mə* is inserted as the exponent of the φ-features corresponding to the

[2] In these varieties, similarly to French, the etymological morphology of the participle is lost. Thus, in the classes preserving the old stressed Thematic Vowel, *-a-* and *-i-*, we find invariable participles such as *la'v-a* 'washed' *cam'm-a* 'called', *dur'm-i* 'slept', *və'n-i* 'come', where the Thematic Vowel creates the adjectival form. In other classes, strong invariable participles occur, e.g. *viawə* 'seen', or, in the case of original… *Vr-tV* sequences, we find the gender alternation of the type of *mɔrə* [masc] vs. *mɔrə-tə* [fem] 'dead', *krəvɛrə* 'covered, masc' and *krəvɛrətə* 'covered, fem', as in (20). This type of alternation, which is attested also in adjectives, e.g. *kiərə* [masc] vs *kiərə-tə* [fem] 'short', in participles regularly appears in stative contexts but, at least for our informant of Celle, it is not excluded in transitive and mid-reflexive contexts where it agrees with the IA. In (20), the element *-t-*, giving rise to the feminine reading, can be now identified with the inflectional exponent.

[3] This solution recalls the one adopted in Roberts (2018), which, however, derives the participle by incorporating v within the Part category and introduces an Aux position including Part.

[4] Roberts (2018) proposes a very similar solution for the OCl-for-SCl phenomena in which the crucial idea is that the OCl is nothing but the instantiation of the φ-features contained in *v*: *v* incorporates with the Part(iciple) head and the OCl with Aux, where it values its features.

IA, so that [R, PPrt] is nothing but R+φ. This holds also in the case of reflexives, where again *v* includes φ agreeing with IA, whereas IA coincides with EA or the latter is excluded.[5] Non-active verbal forms demote the external argument, which can be possibly licensed by means of a specialized preposition/case, or can be simply not realized (Manzini *et al.* 2015).[6] In the progressive contexts introduced by 'stay' exemplified in (13b) and (14), as *mə l əstə faʃaŋŋə* 'I am doing it' (14c), *stay* selects *mə*/*tə* as clitics, behaving like *be*. The problem is that the gerunds can correspond also to transitives and unergatives. We must conclude that *stay* realizes its subject as an IA; the subject of *stay* is coreferential with an argument of the gerund, that, if its IA is independently realized, is the external argument of the gerund.

The selection of the external argument makes recourse to another licenser, i.e. the auxiliary *have* in T, as in (23). This nominal element is merged to *v* and then to T, fixing the person and number agreement of the latter. The past participle retains its nature of modifier of the IA.

(23) CP [<NP, T> [$_{vP}$ NP$_y$[$_v$ have$_y$] [$_{vP}$ [<R$_x$, PPrt/n> (NP$_x$)]]

(21)-(23) allow us to account for the contrast between the unaccusative contexts with *be* in which the subject is the IA agreeing with the participle, and the contexts with lexical unaccusatives. In the contexts with *be*, *mə*(*/tə*) are inserted excluding the 1st/2nd person SCl, while with lexical verbs *dʒə* occurs as the subject clitic. The result is that lexical unaccusatives go together with unergatives and transitives while reflexives behave like predicative constructs. In other words, auxiliaries encode selectional capabilities autonomous with respect to the verb. 3rd person is not involved in this opposition, insofar as shown in section 1.1 it does not differentiate auxiliary and lexical contexts.

In the contexts of lexical verbs, only the agreement of T expresses the referential content of the argument selected by the verb, so requiring the usual SCl form *dʒə*, as in (24). The SCl is merged with the amalgam [T, v] on the basis of its agreeing φ-features. as in (24).

(24) dʒə$_φ$ [T - viŋŋə$_φ$ … [$_{vP}$ [$_R$ vin-$_x$]]] (from (2b))

It is no accident if in deontic constructs in (14d), we find the usual clitic system, with *dʒə*, save for the 1st person clitic of reflexives, where *mə* is inserted, exactly as in the contexts of lexical verbs. Before closing this pf., few words are needed about the 2nd person form *tə* ambiguous for object and subject contexts. We must assume that this is the only lexical entry available for the 2nd person agreement of T or *v*, simply coinciding with the deictic content. Turning now the to the Franco-Provençal dialect of Cantoira, in (18b), we see that in unaccusative and copular constructs with *be* the 1st person excludes the realization of the SCl *dʒi*. Again, agreement with the IA is satisfied by the φ-features realized by the participle/adjective, the same that in the dialects of Celle and Faeto determine the insertion of *mə/tə*.

[5] Not surprisingly, this distribution recalls the contrast between ergative and nominative systems, reducible to the contrast between systems in which *v* licenses the subject (ergative) and systems where it is T that agrees with and licenses the external argument as the subject (D'Alessandro and Roberts 2010).

[6] An anonymous reviewer wonders if it is "possible to assume (on the basis of the data outlined above) that (at least some) unaccusatives are endowed with a little *v* layer". Indeed, we assume that unaccusatives, on a par with other verbal classes, have the layer v (cf. Franco *et al.* 2021). Chomsky (2020a) argues for the proposal that unaccusatives have a "weak" *v*, allowing the IA to be mapped in the C-T domain. This solution fits with our data, whereby the IA is expressed in *v* and in T domains.

4.2 OCls for SCls

Consider the Phasal distribution of clitics in these dialects, as depicted in (25a,b), where in the CP phase SCls are licensed by agreement on the verb in T, as usually in Romance languages. As already suggested for OCls in the vP phase, it is reasonable to assume that SCls are the real heads of agreement in the CP phase (Manzini *et al.* 2020, Savoia *et al.* 2019).

(25) a. CP phase: C SCl OCl T vP

b. vP phase: OCl v [$_{VP}$ R OCl]

In our dialects, when OCls occur in the T domain, the SCls are excluded except for 1st and 2nd person elements, as illustrated in (3). We can see in this partial complementary distribution an effect of DOM. In fact, the 1st/2nd person elements are interpreted independently of the structure of event and can be merged with [R-v], where the OCl agrees with/realizes the φ-features of *v*. Then OCl is merged to [T, R-v], giving rise to (26). If the IA is realized by an OCl merged to T, the procedure terminates, as also suggested by Roberts (2018).[7] We must conclude that 3rd person SCls realize φ-features associated with *v*, with the consequence that only OCls are necessary for the argumental requirements, while SCls realize a property required by T and occur only if φ-features of are not independently expressed.

(26) <OCl$_{\varphi}$, [T, R-v$_{\varphi}$] >

Coming back now to 1st and 2nd person SCls, they are saved and inserted, encoding the EA. We can relate this to the fact that, as being discourse-anchored, they do not involve the φ-features of *v* but are merged directly with T, as suggested in (27) (cf. (25a)).

(27) SCl$_{\varphi}$ [$_{T\varphi}$ OCl R-v]

The surface result is that 1st and 2nd person SCls are obligatorily externalized, while in the other persons the rich agreement inflection is sufficient to express the subject (cf. Chomsky 2015). DOM is involved in the case of the co-occurrence of 2nd person SCl and 1st person OCl, illustrated in (3ii.a), where the order between the clitics is reversed, and we find *mə tə* … in the place of the expected *tə mə*. The reversed order characterizes also contexts with two objects, as in (4b), where *mə* precedes the sequence *tə lu*. The unspecialized nature of *tə* seems to be at the root of the phenomenon, to the effect that *mə* is merged to the usual structure in which *tə* combines with T-v, possibly already merged with *lu*, <mə, [tə, [(lu-)T-vφ]>. The sequence of pronominal exponents externalizes the order 1st – 2nd revealing the interpretive role of the referential hierarchy discussed by Kiparsky (2008; cf. fn. 1).

In many Romance languages with SCls, agreement relations in simple contexts include the agreement exponent of the verb and the SCl, understood as a sort of a discontinuous exponent of the T head. Nevertheless, in the presence of OCls we have at least two possibilities (in relation to the person), namely SCls combine with OCls or only OCls occur (Manzini and Savoia 2005). Tentatively, we can suggest the generalization on the distribution of agreement in (28) (cf. Roberts 2018).

[7] Roberts (2018: 261) explains the OCl-for-SCl distribution in the Valdôtain and Piedmontese dialects, through two steps: clitics are generated by a process of morphological fission (Halle and Marantz 1996), " 'splitting off' of (a subset of) features of a head to form a separate morpheme"; in OCl-for-SCl dialects the fission process is applied to Aux-level. The lack of a rule fusing OCls and SCls excludes the realization of the SCl under Aux. Interestingly, our approach implies a similar conceptualization, based on the role of the OCl as mapping the φ-features of Aux.

(28) Agreement generalization: Agr T a. introduces a specialized sub-set of φ-features
b. copies the set of φ-features associated with *v*

(28a) introduces SCls as a sort of duplicate of the agreement on the verbal head, as generally in Northern Italian dialects. (28b) corresponds to the IM of φ-features of *v* as inflectional morphemes (unaccusatives, passives, reflexives) and possibly OCl. In our dialects, (28b) is generally applied. In dialects such as that of Cantoira in (18), (28b) is applied except for the 1st person SCl combining with the 2nd person OCl. We conclude that in the particular context the necessary properties of the deictic elements have a full realization.

Finally, consider the other clitics. The 2nd person OCl and SCl share the same forms *tə* and *və*, while the 1st person plural is *sə*, i.e. the reflexive element. The reflexive element *sə* occurs in all the persons except for 1st and 2nd singulars, in the reflexive paradigm, cf. (5), where it is in complementary distribution with all SCls. We follow the proposal of Manzini *et al.* (2016), whereby *sə* satisfies the IA but its content is that of a variable, that finds its referential properties in the agreement features of the verb, its antecedent. In the case of *mə* and *tə* the reflexive interpretation derives from their coincidence with the verbal inflection-the subject. In other words, the exclusion of SCls can be traced back to the same mechanism working in the case of accusative clitics, except for 1st and 2nd person contexts. An anonymous reviewer raises the question why "this isn't possible with other pronouns. Or conversely, why the specialised 'se' form cannot extend to 1/2P". Indeed, the generalized occurrence of 'se' in reflexive for all persons is attested in Northern Italian and Rhaeto-Romance varieties where yet the clitic of 1st person tends to be however inserted (cf. Manzini and Savoia 2005: §4.3). The insertion of the element *se/si* introducing the variable is generally associated with the nominal φ-features (D) of the subject, including also definiteness/specificity, able to fix the denotation of *se/si*. Along these lines, we can suggest that such nominal properties work as the lexical restriction of *se*, excluding the coreference with 1st/2nd person elements. This can explain why 1st and 2nd plural persons, insofar as encoding also the 3rd person reference, have the reflexive *se* in many dialects, including the Celle and Faeto ones.

A final point concerns the special 1st person form ɛ of the auxiliary/copula in (14)-(15) for Faeto. In this variety the paradigms of the present of *be* and *have* are distinct except for the 1st person, where only the form ɛ occurs independently of the nature of the predicate. Systems with totally or partially overlapping paradigms of auxiliaries are discussed in Manzini and Savoia (2011), which propose to identify them with forms of *be*. This analysis suggests that in the case of agentive verbs *be* is obliged to select the EA of the lexical verb, thus excluding the insertion of *mə*.

5. Final remarks

In this article we started from two main points: the syntax of auxiliaries and the occurrence of clitics in Apulian Franco-Provençal. In this language object and subject clitics are in a partially complementary distribution related to DOM constraints and show an unexpected occurrence of the object clitics *mə, tə* instead of the SCl in copular and non-active contexts. The coincidence between *be* as auxiliary and copula provides evidence in favor of the analysis of auxiliaries as verbs in the proper sense, embedding a predicative clause. The participle has nominal inflectional properties associated with the argument organization of unaccusatives and passives. Our analysis is based on the recent formalization of modification structure proposed in Chomsky (2020a, b), where units of agreeing lexical elements are created by (pair-)merge. This framework makes a more natural and interesting treatment of cliticization possible; in our case, *mə/ tə* are the exponents of 1st/2nd person IA, which are merged to *v* and then to T, interpreting the agreement of the verb instead of SCl.

References

Baldi, Benedetta, and Leonardo M. Savoia. 2019. "Auxiliary Selection in Italian Dialects: Person Split, OCls and *Raddoppiamento Fonosintattico.*" *Quaderni di Linguistica e Studi Orientali* QULSO, 5: 55-105.

Baldi, Benedetta, and Leonardo M. Savoia. 2021. "Possessives, from Franco-Provençal and Occitan Systems to Contact Dialects in Apulia and Calabria." *Languages*, 6 (63): 1-28.

Belvin, Robert, and Marcel den Dikken. 1997. "There, *Happens*, to, Be, Have." *Lingua*, 101 (3-4): 151-183.

Belletti, Adriana. 1990. *Generalized Verb Movement*. Torino: Rosenberg and Sellier.

Belletti, Adriana. 2005. "(Past-)participle Agreement." In *Blackwell Companion to Syntax*, vol. III, ed. by Martin Everaert, and Henk C. van Riemsdijk, 493-521. Malden, MA: Blackwell.

Bentley, Delia, and Thórhallur, Eythórsson. 2004. "Auxiliary Selection and the Semantics of Unaccusativity." *Lingua* 114 (4): 447-471.

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

Chierchia, Gennaro. 2004. "A Semantics for Unaccusatives and its Syntactic Consequences." In *The Unaccusativity Puzzle. Explorations of the Syntax-Lexicon Interface*, ed. by Artemis Alexiadou, Elena Anagnostopoulou and Martin Everaert, 22-59. Oxford: Oxford UP.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: A Life in Language,* ed. by Michael Kenstowicz, 1-54. Cambridge, MA-London: MIT Press.

Chomsky, Noam. 2015. "Problems of Projections." In *Structures, Strategies and Beyond. Studies in honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, Simona Matteini, 3-16. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company.

Chomsky, Noam. 2020a. *The UCLA Lectures.* (April 29-May 2, 2019). <https://ling.auf.net/lingbuzz/005485> (06/2021).

Chomsky, Noam. 2020b. "Puzzles About Phases." In *Linguistic Variation: Structure and Interpretation: Contributions in Honor of M. Rita Manzini*, ed. by Ludovico Franco and Paolo Lorusso, 163-167. Berlin: De Gruyter.

Chomsky, Noam, Ángel Gallego, and Dennis Ott. 2019. "Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges." In *Generative Syntax: Questions, Crossroads, and Challenges*, *Catalan Journal of Linguistics*. *Special Issue*, ed. by Ángel Gallego, and Dennis Ott, 229-261.

D'Alessandro, Roberta, and Ian Roberts. 2010. "Past Participle Agreement in Abruzzese: Split Auxiliary Selection and the Null-Subject Parameter." *Natural Language and Linguistic Theory*, 28 (1): 41-72.

Dixon, Robert M.W. 1994. *Ergativity*. Cambridge, MA-New York, NY: Cambridge UP.

Fillmore, Charles. J. 1968. "The Case for Case." In *Universals in Linguistic Theory*, 1-88, ed. by Emmon Bach, and Robert T. Harms. New York, NY: Holt, Rinehart, and Winston.

Franco, Ludovico, M. Rita Manzini, and Leonardo M. Savoia. 2021. "Locative PS as General Relators. Location, Direction, DOM in Romance." In *Into Adpositions. New Formal Perspectives on the Structure of the PP and its Variation*, *Linguistic Variations, Special Issue.* 135-173, ed. by Víctor Acedo-Matellán, Theresa Biberauer, Jaume Mateu, and Anna Pineda. Amsterdam: Benjamins.

Halle, Morris, and Alec Marantz. 1996. "Distributed Morphology and the Pieces of Inflection." In *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Kenneth L. Hale and Samuel Jay Keyser, 111-176. Cambridge, MA: MIT Press.

Kayne, Richard S. 1993. "Toward A Modular Theory Of Auxiliary Selection." *Studia Linguistica*, 47 (1): 3-31.

Kiparsky, Paul. 2008. "Universals Constrain Change; Change Results in Typological Generalizations." In *Linguistic Universals and Language Change*, ed. by Jeff Good, 23-53. Oxford: Oxford UP.

Manzini, M. Rita, and Ludovico Franco. 2016. "Goal and DOM Datives." *Natural Language and Linguistic Theory*, 34 (1): 197-240.

Manzini, M. Rita, and Leonardo M. Savoia. 2005. *I Dialetti Italiani e Romanci*, vol. III. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita, and Leonardo M. Savoia. 2010. "Syncretism and Suppletion in Clitic Systems: Underspecification, Silent Clitics or Neither?" In *Syntactic Variation. The Dialects of Italy*, ed. by Roberta D'Alessandro, Adam Ledgeway, and Ian Roberts, 86-101. Cambridge: Cambridge UP.

Manzini, M. Rita, and Leonardo M. Savoia. 2011. *Grammatical Categories: Variation in Romance Languages*. Cambridge, MA: Cambridge UP.

Manzini, M. Rita, and Leonardo M. Savoia. 2014. "Linkers in Aromanian in Comparison to Albanian and Romanian." *Rivista di Grammatica Generativa*, 36: 83-104.

Manzini, M. Rita, and Leonardo, M. Savoia. 2017. "Gender, Number and Inflectional Class in Romance: Feminine/Plural -a." (Olomouc, Czech Republic, 7-9 June 2018). In *Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2018*, ed. by Joseph Emonds, Markéta Janebová and Ludmila Veselovská, 263-282. Olomouc: Palackỳ University Olomouc.

Manzini, M. Rita, Anna Roussou, and Leonardo M. Savoia. 2016. "Middle-Passive Voice in Albanian and Greek." *Journal of Linguistics*, 52 (1): 111-150.

Manzini, M. Rita, Leonardo M. Savoia, and Benedetta Baldi. 2020. "Microvariation and Macrocategories: Differential Plural Marking and Phase theory." *L'Italia Dialettale*, vol. LXXXI: 189-212.

Manzini, M. Rita, Leonardo M. Savoia, and Ludovico Franco. 2015. "Ergative Case, Aspect and Person Splits: Two Case Studies." *Acta Linguistica Hungarica*, 62 (3): 297-351.

Manzini, M. Rita, Leonardo M. Savoia, and Ludovico Franco. 2020. "DOM and Dative in (Italo-) Romance." In *Case, Agreement, and their Interactions: New Perspective on Different Argument Making*, ed. by András Bárány and Laura Kalin, 219-268. Berlin: De Gruyter.

Moro, Andrea. 1997. *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*. Cambridge: Cambridge UP.

Roberts, Ian G. 1993. "The Nature of Subject Clitics in Francoprovençal Valdotain." In *Syntactic Theory and the Dialects of Italy,* ed. by Adriana Belletti, 319-353. Torino: Rosenberg et Sellier.

Roberts, Ian G. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.

Roberts, Ian G. 2018. "Object Clitics for Subject Clitics in Francoprovençal and Piedmontese." In *Structuring Variation in Romance Linguistics and Beyond*, ed. by Mirko Grimaldi, Rosangela Lai, Ludovico Franco, and Benedetta Baldi, 257-267. Amsterdam: Benjamins.

Savoia, Leonardo M., and M. Rita Manzini. 2010. "Lexicalization of 3rd Person Object Clitics: Clitic Enclisis and Clitic Drop." In *Syntactic Variation. The Dialects of Italy*, ed. by Roberta D'Alessandro, Adam Ledgeway, and Ian G. Roberts, 102-118, Cambridge-New York, NY: Cambridge UP.

Savoia, Leonardo M., Benedetta Baldi, and M. Rita Manzini. 2019. "Asymmetries in Plural Agreement in DPs." (Olomouc, Czech Republic, 7-9 June 2018). In *Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2018*, ed. by Joseph Emonds, Markéta Janebová, and Ludmila Veselovská, 204-224. Olomouc: Palacký University Olomouc.

Savoia, Leonardo M., and M. Rita Manzini. 2010. "Les clitiques sujets dans les variétés occitanes et francoprovençales italiennes." *CORPUS*, 9: 165-189.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: E. Di Domenico (2021) Bilingualism and (outward) over-explicitness in the choice of subject anaphoric devices. *Qulso* 7: pp. 51-73. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12002

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Bilingualism and (outward) over-explicitness in the choice of subject anaphoric devices*

Elisa Di Domenico
Università per Stranieri di Perugia (<elisa.didomenico@unistrapg.it>)

Abstract:

In this work we propose an explanation (the Decreased Activation Hypothesis) for some *prima facie* conflicting findings concerning the choice/interpretation of subject anaphoric devices in late bilinguals (LBs). While some studies report an overuse of overt pronouns (often claimed to be a default form employed by LBs), other studies report an overuse of lexical DPs (interpreted as a sign of LBs' over-explicitness). Our proposal is that over-explicitness in bilinguals is just outward and the use of seemingly over-explicit forms (overt pronouns or lexical DPs) stems from LBs difficulty to cope with referents' activation. Then, starting from the observation that whenever overuse of overt pronouns is reported a null subject language is at least involved, and whenever overuse of lexical DPs is reported a non-null subject language is at least involved, we explore the way in which subject anaphoric devices are employed in (some) null subject languages and in (some) non-null subject languages, finally arguing that LBs of a null and a non-null subject language may choose to be seemingly overexplicit in two different ways.

Keywords: *Accessibility, Bilingualism, Over-Explicitness, Subject Anaphoric Devices*

1. Introduction

A recent discovery in the field of bilingualism is that bilinguals (in particular late bilinguals, henceforth LBs, i.e. adult speakers who have acquired a second language after puberty) tend to overuse overt subject pronouns in their null subject language, regardless of the fact that they have correctly set the relevant parameter and that their other language is also a null subject language. In specific, overt subject pronouns are used by LBs also in topic continuity and interpreted also as co-referent with

* Thanks to Alexander Cairncross, Anna Cardinaletti, Luigi Rizzi, Ludovica Serratrice and two anonymous reviewers.

a topical antecedent at a higher rate with respect to native speakers of a null subject language, which prefer a null pronoun to pick up a topical antecedent.

An influential explanation for this finding is the Interface Hypothesis (Sorace and Filiaci 2006; Tsimpli, Ianthi, and Sorace 2006; Sorace 2011) which, in its more recent version, states that properties at the syntax-discourse interface are particularly vulnerable, and that bilinguals have more limited processing resources with respect to monolinguals: this makes it difficult for them to compute topicality. Overt pronouns thus represent a convenient 'default' strategy, in that, through their phi-features, they help identifying the intended antecedent, while, at the same time, they allow to avoid lexical retrieval, which may be costly for these speakers (Sorace 2011, 2016 a.o.). Another proposed explanation is the Pragmatic Principles Violation Hypothesis (Lozano 2016) which assumes that LBs will violate more often pragmatic principles banning redundancy than pragmatic principles banning ambiguity.

Other studies report that LBs overuse lexical DPs with respect to monolinguals, a data which does not seem to support the first explanation, but rather the view that LBs are "over-explicit" with respect to monolinguals (Ryan 2015 a. o.).

The question, however, is why LBs are over-explicit, and why over-explicitness is instantiated in two different ways: overuse (and interpretation also in co-reference with a subject antecedent) of overt pronouns, and overuse of lexical DPs.

On the basis of production data from LBs of two null subject languages, Di Domenico, Baroncini and Capotorti (2020) proposed that over-explicitness might be just outward, and that the specific problem that these speakers are faced with is in establishing the prominence of an antecedent (topicality) when more than one referent is active, because the presence of an additional character decreases referents' activation. If so, overt pronouns are not a default option for LBs of two null subject languages, but rather the sufficiently explicit device to employ in this condition.

In this work we will explore the possibility that what we will call the "Decreased Activation Hypothesis" can be extended to explain also the other data observed in the relevant literature. Starting from the observation/generalization that whenever overuse of overt pronouns is reported, a null subject language is at least involved, and whenever overuse of lexical DPs is reported a non-null subject language is at least involved, we will finally propose that there are two ways of being "seemingly over-explicit" (the "null subject language centered" strategy and the "non-null subject language centered" strategy) and that LBs of a null and a non-null subject language, given their bilingual competence, may choose one or the other. This requires on one side a comparative analysis of how anaphoric devices are employed in null and non-null subject languages, and on the other the assumption that bilingual speakers may choose among the options offered by their multilingual competence.

In Section 2 we will review some relevant literature concerning the use and interpretation of subject anaphoric devices by LBs. As we said, some studies report an overuse of overt pronouns while other studies report an overuse of lexical DPs. The baseline on which 'overuse' is calculated is usually represented by native, monolingual speakers.[1] In Section 3 we will explore how subject anaphoric devices are chosen and interpreted by native monolingual speakers of null and non-null subject languages. In Section 4 we will outline the core of our proposal, examining as well some cases which require an additional explanation. In Section 5 we will draw some conclusions.

[1] An exception is Torregrossa and Bongartz (2018). We use the term "monolingual speakers" to refer to speakers who have grown with one language, which is the language they currently employ in everyday life. Of course a monolingual speaker may have (and this is generally the case) some knowledge of another language acquired later in life.

2. The use and interpretation of subject anaphoric devices by (Late) Bilinguals: a review of some relevant literature

2.1 The overuse and over-acceptance of overt pronouns

Tsimpli *et al.* (2004) examined a group of 20 adult speakers with Italian as a first language (L1) who had a minimum of 6 years of residence in Britain, and were hence very advanced (near-native, according to White and Genesee's 1996 criteria) speakers of English as a second language (L2), but still used their L1 on a regular basis. They found significant differences between this experimental group and a control group of 20 monolingual speakers of Italian in the interpretation of overt pronouns. [2] They used experimental bi-clausal sentences like those in (1), with a null or an overt pronoun in the temporal clause and two possible antecedents in the main clause. The temporal clause could either precede (1.a and 1.b) or follow (1.c and 1.d) the main clause:

(1) a. Quando *pro* attraversa la strada, l' anziana signora saluta la ragazza
when *pro* crosses the street the old lady greets the girl
b. Quando lei attraversa la strada, l' anziana signora saluta la ragazza
when she crosses the street the old lady greets the girl
c. L' anziana signora saluta la ragazza quando *pro* attraversa la strada
the old lady greets the girl when *pro* crosses the street
d. L' anziana signora saluta la ragazza quando lei attraversa la strada
the old lady greets the girl when she crosses the street

The experimental group showed a significantly greater tendency to interpret the overt pronoun as co-referent with the subject of the main clause, when compared to the control group. [3] No significant differences between the experimental and the control group were found in production, elicited through a story telling task. These differences in interpretation were assumed to be an effect of attrition on the L1 caused by the L2.

The authors distinguish between the setting of the null subject parameter on one side, and the way in which the options made available by the grammar (i.e. in null subject languages, the availability of null subjects and of the postverbal subject position) are assigned interpretable features relevant at the LF-interface, predicting that attrition will affect the latter only.[4] As an effect of attrition, an interpretable feature that is specified in the L1 will become unspecified due to the absence of a similar interpretable feature in the L2 in the same syntactic context.

[2] These data were part of a more comprehensive study concerning the syntax of subjects in Greek and Italian speakers of L2 English living in Britain. The Greek experimental group did not significantly differ from the Greek monolingual controls in the interpretation and production of overt pronouns. They differed from controls in producing less postverbal subjects, and in a more ambiguous interpretation of preverbal and postverbal indefinite subject.

[3] Interpretations were elicited through a picture verification task. Participants had to choose one among three pictures as the antecedent of the pronoun (either the picture showing an old lady, or the girl, or someone else).

[4] This entails, according to the authors, that the interpretation of null subjects should not be affected by attrition, since the availability of null subjects is regulated by a parametric choice The authors also found, however, a significantly higher subject oriented interpretation of the null pronoun in sentences like (1.c) by LBs: an unexpected result, that they explain hypothesizing that the experimental group may treat the subordinate clause as a non-finite one (as in the English subject-control sentence *The old woman greets the girl when crossing the street*).

The results concerning the interpretation of overt pronouns were replicated by Sorace and Filiaci (2006), and by Belletti, Bennati and Sorace (2007), who studied L2 speakers of Italian with L1 English, with experimental materials similar to those employed by Tsimpli *et al.* (2004). While overt pronouns were significantly more interpreted in co-reference with the subject antecedent by the experimental group, no significant differences between the two groups were observed in the interpretation of null subjects.[5] Sorace and Filiaci (2006) first formulated the Interface Hypothesis, which states that properties at the syntax-discourse interface are more vulnerable than purely syntactic properties and subject to protracted optionality even in near-native speakers. Belletti, Bennati and Sorace (2007) also employed a story telling task which revealed significant differences between the experimental and the control group in production: the rate of overt pronouns in the experimental group (14%) was significantly higher than in the control group (4%). No significant differences were found for what concerns null pronouns or lexical DPs. One possibility to explain the data, as proposed by Belletti, Bennati and Sorace 2007, is that the strong (in the sense of Cardinaletti and Starke 1999) overt pronoun of Italian is re-analysed as weak, possibly on the basis of the L1 option, and hence used in the place of *pro* (the null analogue of the weak overt pronoun).[6]

The authors also mention (see their Footnote 17) an alternative possible analysis, according to which overt pronouns might be a "default" option that speakers resort to when they are faced with "different sorts of processing difficulties" (Belletti, Bennati and Sorace 2007: 673).

Suggestive evidence for this alternative analysis, pursued in greater details by Sorace (2011, 2016), could be provided by growing evidence showing that overt subject pronouns are overused also by LBs who speak two null subject languages.

Margaza and Bel (2006) administered a cloze task to intermediate and advanced learners of L2 Spanish with L1 Greek. The task included 40 empty subject positions (in matrix and embedded clauses) which participants had to fill. Results revealed significant differences between the intermediate group and the control group in matrix clauses, where intermediate learners admitted null subjects less often than the control group. No significant differences were observed, however, between the advanced group and the control group.

Learners of L2 Spanish with L1 Greek are also studied by Lozano (2006, 2018). Lozano (2018) administered an acceptability judgment task to intermediate and advanced (lower advanced and upper advanced) learners and a control group. L2 learners accepted overt pronouns at a significantly higher rate with respect to controls in topic continuity contexts.[7] In topic shift

[5] Some interpretation differences were observed between (1.b) and (1.d) in the control group, and between (1.a) and (1.c) in both groups. We will deal with these differences in Section 3.

[6] This analysis might be reinforced, as the authors argue, by a weak use of the overt pronoun attested in Italian (Cardinaletti 2004a, 2004b; Frascarelli 2007). See Section 3, ex. (14) in particular. The strong/weak distinction of personal pronouns (Cardinaletti and Starke 1999) is based on a number of prosodic, morphological, syntactic, semantic and referential properties. From the referential point of view, weak pronouns are assumed to have a prominent antecedent, while strong pronouns a non-prominent one. Coherently, a null pronoun is a weak pronoun. Overt third person pronouns of the *lui/lei/loro* kind in Italian are usually strong pronouns, while *egli/ella/esso/essa* are weak pronouns. Since the latter are disappearing in colloquial Italian, a weak form of *lui/lei/loro* is assumed to be developing in current Italian. English personal pronouns are assumed to have a weak and a strong form which are not morphologically distinct (Cardinaletti 2004b: 133). An exception is 'it' which is only weak: it cannot be coordinated, modified, focalized; as an object pronoun, it must precede the particle in particle verb constructions (*They turned on it, Cardinaletti 2004b ex. (15.a)), while as a subject pronoun it cannot be separated from the verb by a parenthetical (*It, I think, is fine).

[7] Following Lambrecht (1994: 118) a topic is "the thing which the proposition expressed by the sentence is about". It is very often the case that topic and subject coincide, although the two notions should be kept separate.

contexts, however, L2ers did not significantly differ from controls: all groups accepted overt pronouns and did not accept null pronouns. While underlining the fact that the L1-L2 similarity does not facilitate the task, the author interprets these results in terms of the Pragmatic Principles Violation Hypothesis (Lozano 2016), according to which L2ers are more inclined to violate pragmatic principles banning redundancy than pragmatic principles banning ambiguity.

Di Domenico and Baroncini (2019) examined the productions of L1 Greek- L2 Italian speakers through a story telling task analogous to the one used by Tsimpli *et al.* (2004) and Belletti, Bennati and Sorace (2007). They had three different groups of bilinguals, and two control groups: bilinguals from birth living in Greece, bilinguals from birth living in Italy, late bilinguals living in Italy, Italian natives and Greek natives. LBs (as well as the two groups of bilinguals from birth) reached a near-native level of proficiency in Italian. The authors found that LBs used overt subject pronouns at a significantly higher rate with respect to all the other groups, while they did not differ in the amount of null subjects or lexical DPs employed. This suggests, as the authors acknowledge, that the overuse of overt pronouns characterizes LBs (and not bilinguals from birth) independently from cross-linguistic influence. Interestingly, the rate of overt subject pronouns used by late bilinguals in this study (14.50%) is similar to the one reported by Belletti, Bennati and Sorace (2007).

Taken together, these three studies support the view that cross-linguistic influence may not be involved in (or may not be the only cause of) the observed overuse (and over-acceptance) of overt subject pronouns by LBs.[8]

Sorace (2011, 2016) has suggested that computing topicality might be costly for LBs whose processing resources are more taxed (Clahsen and Felser 2006; Abutalebi 2008; Kroll and Bialystock 2013). Overt pronouns might be thought of as a convenient default option which LBs may resort to: with respect to null pronouns they may be unambiguous, since they have phi-features, while, at the same time, they allow the avoidance of lexical retrieval, which may also be costly for LBs.

Contemori and Dussias (2016) interpret their data along these lines. They investigated the choice of referring expressions in English by a group of highly proficient L1 Spanish- L2 English speakers, compared to a group of English monolingual speakers. They used an adaptation of Arnold and Griffin's (2007) task (Experiment 1), and of Hendriks, Koster and Hoeks (2014) task (Experiment 2). The first task is meant to elicit subject anaphoric devices in contexts of topic maintenance (with one referent, two referents of the same gender, two referents differing in gender) while in the second task there are contexts in which the topic is maintained and also contexts in which the topic changes (see also Section 3).

Experiment 1 reveals a significantly higher use of overt pronouns (and a lesser use of lexical DPs) by L2ers in all conditions, particularly significant in the condition with two referents differing for gender. Experiment 2 reveals marginally significant differences in the same direction: L2ers used less lexical DPs than native speakers in the conditions where the topic is maintained. No significant differences between the experimental and the control group were observed in topic shift contexts. The authors interpret their results assuming that L2ers choose overt pronouns because they are a default form which is easy to select and to produce. If there were cross-linguistic influence effects, the authors argue, the data would reveal a high rate of null pronouns, contrary to fact.

See Rizzi (2005, 2018) as well as Section 3 below for some discussion.

[8] See also Bini (1993) for L1 Spanish- L2 Italian LBs and Judy (2015) for L1 Farsi- L2 Spanish LBs.

2.2 The overuse of lexical DPs

While Contemori and Dussias (2016) have found an overuse of overt pronouns by L2ers of a non-null subject language with a null subject L1, some studies report an overuse of lexical DPs by speakers with an analogous combination of languages.

Ryan (2015) analyses the use of referring expressions by a group of Chinese learners of English as well as by a control group of native speakers, through a film retelling task. As the author notes, evidence has been provided in the literature that second language learners appear to violate a principle of economy by producing anaphoric forms that are over-explicit, such as lexical DPs where pronouns or null pronouns ("zero anaphora", or "zero" in his terms) would be expected. The initial hypothesis investigated is whether this claimed over-explicitness might be simply a by-product of the use of different narrative strategies by second language learners, and hence the higher rate of more explicit devices in L2ers' narrations is due to the fact that less accessible referents are taken up. The author thus develops, on the basis of Toole (1996), a coding system to determine referents accessibility, estimated as the sum of multiple weighted factors (grammatical role, presence in main or embedded clauses, animacy, parallel function, main vs. secondary character). Differences are revealed in this respect between native speakers and L2ers, the latter referring to most accessible antecedents significantly more than native speakers, and to antecedents at a medium degree of accessibility significantly more than native speakers. When analysing referring expressions, however, it appeared that L2ers' use of high accessibility markers (particularly pronouns) was far lower than what could be accounted for by the observed differences in referents' accessibility, suggesting that over-explicitness did characterize these data. Ryan proposes that L2ers over-explicitness is in turn motivated by a concern for communicative clarity.

Chini (2005) analyses the productions in Italian of a group of 8 adult advanced learners of Italian with L1 German and of a control group of 13 native speakers of Italian, collected through a film retelling task. The advanced learners were also asked to produce a summary in their L1, i.e. German. She analyses the use of different referential forms in referents' introduction and reference maintenance contexts. [9] No differences were found in the referential forms employed for referents introduction: in all the three corpora, indefinite lexical DPs (or a few proper names for the protagonist) were used. As far as reference maintenance is concerned, the author distinguishes three situations: a) the reference maintaining devices which are employed just after the introduction of a referent (one or two sentences after) analysed in the three corpora; b) reference maintenance contexts as a whole (L2 and L1 Italian corpora); c) "anaphoric chains" (L2 and L1 Italian corpora), i.e. clause sequences where one and the same referent is addressed. As for the first situation, she found that lexical DPs are used by L2ers and L1 German speakers at a higher rate with respect to L1 Italian speakers, namely 44% by L2ers, 42% by L1 German and only 17% by L1 Italian speakers. Null pronouns are used by L2ers (18%) and by L1 Italian (12%), but almost never, as expected, by L1 German speakers. As for the (quite surprising) higher use of null pronouns by L2ers with respect to L1 Italian speakers, the author argues that null pronouns in L2ers appear to be used in the place of relative (or pseudo-relative) pronouns, which are very frequent in Italian but less so in L2ers.[10] In reference maintenance

[9] Shifting of reference contexts are coded in the corpus, but unfortunately not counted or analyzed in the paper. The paper considers both reference to subjects and reference to objects, but we will be mainly concerned with reference to subjects, as in the rest of this work.

[10] The example in (i) is produced by an L2er, while an L1 Italian would produce (ii) (Chini 2005: 79):

contexts as a whole, definite lexical DPs were used four times more often in L2ers than in Italian natives. Overt pronouns are also used three times more by L2ers than by Italian natives, and also in contexts of high referential continuity. As the author argues, "this seems to be a typical co-referential choice for (these German) learners" (Chini 2005: 85). The use of null pronouns is lower in L2ers (22%) than in native speakers (34%). In anaphoric chains, a consistent use of null pronouns is observed in both L1(74%) and L2 (69%) Italian speakers. Lexical DPs (here proper names only) are very few in both groups. A remarkable difference is observed instead for what concerns overt pronouns (19% L2ers, 7% L1 Italian).[11] The author assumes that the over-explicitness found in L2ers is due in part to a source language influence. She observes, however, that most of the over-explicit expressions are found (at least in some narratives, no quantitative evidence is given) in transitions from main clauses to dependent clauses and at the beginning of scenes and sub-episodes, i.e. in contexts of syntactic and/or textual discontinuity which may discourage the use of implicit means.

Torregrossa and Bongartz (2018) analyse, through a story telling task, the production of referring expressions in Italian by 17 German-Italian bilingual adolescents (mean age 13 y. o., range 11.9-14.1) living in Germany. Their aim is precisely to identify over-explicit (over-specific in their terms) uses of referring expressions. In order to establish a base-line, instead of the usual comparison with a control group, they assess the accessibility of a referent considering the grammatical role of the antecedent (subject vs. non-subject) and the presence of intervening referents (competition).

They then focus on those contexts in which an overt pronoun or a lexical DP are used when a more reduced form (a null pronoun or a clitic) would be appropriate. They finally correlate their findings to an index of dominance, the Bilingual Index Score, for each participant (a negative number indicates dominance in German, while a positive one dominance in Italian). Considering the total amount of over-explicit forms in subject and object position (*N*=34), they found that the number of over-explicit lexical DPs (*N*=29) greatly exceeded ($p<.001$) the amount of overt pronouns (*N*=5). Furthermore, they found that the lower the Bilingual Index Score of the participants (i.e. the more dominant in German they were), the more over-explicit forms they produced in Italian. Considering the type of over-explicit forms mostly employed by their participants, the authors argue that the use of lexical DPs cannot be motivated by cross-linguistic influence, since lexical DPs are attested in both languages. Their findings, they suggest, are rather compatible with a processing account of bilinguals' difficulties with referring expressions. This study brings to light the fact that even simultaneous bilinguals may be over-explicit, to a certain extent.[12] It also indicates that over-explicitness inversely correlates

(i) C'è Charlie Chaplin e trova una bandiera
there is Charlie Chaplin and finds a flag
(ii) C'è Charlie Chaplin che trova una bandiera
there is Charlie Chaplin who finds a flag

[11] Clitics are also used differently in the two groups of speakers: 4% in L2ers, 16% in L1 Italian. Clitics are the second major device used by L1 Italian speakers, but the least used by L2ers. Since there is however no increase of lexical DPs in L2ers, we assume that lesser use of clitics entails that some overt pronouns might be used for object reference as well, and also that some null objects should be present in L2ers narratives.

[12] The authors do not give indications that concern the rate of over-explicit forms with respect to the total of referring expressions. Their data set consists of 298 referring expressions (p. 13), plus 7 cases of clitic omission (p.17). This would give a percentage of over-explicit forms of around 11.15%. In the data set, however, forms used for referents introduction are included, and their amount is not specified. This means that the rate of over-explicit forms is certainly higher than 11.15%.

with dominance: this result is particularly interesting assuming that dominance is associated with lesser processing difficulties.

Leclercq and Lenart (2013) analyse the anaphoric forms used by different groups of child and adult speakers acquiring French and English while retelling a short film (a 5m. cartoon with two protagonists, a dog and a boy). While children tended to use under-explicit forms when compared to adult native speakers, adult L2 speakers of French and English were over-explicit. Here we will briefly report on the results that concern only subject anaphoric devices used by adult native speakers of French and English, and advanced L2 adult speakers (L1 French-L2 English; L1 English-L2 French).[13] Although French and English are two non-null subject languages, Leclercq and Lenart's (2013) data reveal that null pronouns (zero forms in their terms) are quite attested in the L1 English speakers' productions (21.35%): coordinated structures are quite frequently used when reference is maintained to the same topic. This is not so in the L1 French speakers' productions, where null pronouns amount to 4.35%. Interestingly, this characteristic is not transferred in L2 French, where the rate of null pronouns is 1%: the authors suggest that either these speakers are sensitive to the property of the L2, or that they disfavour the null pronoun option due to their tendency to be over-explicit. L2 productions indeed generally display a higher proportion of lexical DPs in both French and English when compared to the productions of L1 speakers. In L2 English, where we observe the highest proportion of lexical DPs, overt pronouns are reduced with respect to the productions in the other groups. The authors interpret over-explicitness as a means to avoid ambiguity.[14] As they underline, over-explicitness is attested also in advanced learners, which suggests "a deeply ingrained tendency to overspecification among learners" (Leclercq and Lenart 2013: 27).

2.3 Some discussion and an alternative proposal

In Table 1 we briefly summarise the results of the studies reviewed so far.

	STUDY	LANGUAGE(S) INVESTIGATED	LANGUAGES OF THE PARTICIPANTS	MODALITY	OVER-EXPLICIT SUBJECT ANAPHORIC DEVICE
1	Tsimpli *et al.* 2004	Italian	L1 Italian L2 English	Interpretation Production	Overt pronouns
2	Sorace and Filiaci 2006	Italian	L1 English L2 Italian	Interpretation	Overt pronouns
3	Belletti, Bennati and Sorace 2007	Italian	L1 English L2 Italian	Interpretation Production	Overt pronouns
4	Margaza and Bel 2006	Spanish	L1 Greek L2 Spanish	Production	Overt pronouns

[13] Leclercq and Lenart (2013) had 10 experimental groups (with 6 to 10 participants in each group): L1 French children (4, 7, 10 y. o.), L1 English children (4,7,10 y. o.), L1 French-L2 English adults (intermediate; advanced), L1 English-L2 French adults (intermediate; advanced). There were also two control groups: adult L1 English speakers and adult L1 French speakers.

[14] Under-explicitness in children is interpreted as reflecting children's difficulty in taking into account the addressee's perspective.

5	Lozano 2018	Spanish	L1 Greek L2 Spanish	Acceptability	Overt pronouns
6	Di Domenico and Baroncini 2019	Italian (Greek)	L1 Greek L2 Italian	Production	Overt pronouns
7	Contemori and Dussias 2016	English	L1 Spanish L2 English	Production	Overt pronouns
8	Ryan 2015	English	L1 Chinese L2 English	Production	Lexical DPs
9	Chini 2005	Italian (German)	L1 German L2 Italian	Production	Overt pronouns Lexical DPs
10	Torregrossa and Bongartz 2018	Italian	German- Italian 2L1	Production	Lexical DPs (overt pronouns)
11	Leclercq and Lenart 2013	English French	L1 French- L2 English; L1 English- L2 French	Production	Lexical DPs

Table 1 – Summary of the results of the reviewed studies

The studies reviewed so far, thus, report that LBs (and in one case, i.e. Torregrossa and Bongatrz, 2018, also unbalanced bilinguals from birth) differ from native speakers in their choice/interpretation of subject anaphoric devices in different ways. In specific: a) LBs overuse overt pronouns and/or interpret/accept overt pronouns as co-referent with a topical antecedent in their null subject L2 when their L1 is a non-null subject language (Tsimpli *et al.* 2004; Chini 2005; Sorace and Filiaci 2006; Belletti, Bennati and Sorace 2007) but also when their L1 is a null subject language (Margaza and Bel 2006; Lozano 2018; Di Domenico and Baroncini 2019). b) LBs overuse overt pronouns in their non-null subject L2 when their L1 is a null subject language (Contemori and Doussias 2016). c) LBs overuse lexical DPs in their null subject L2 when their L1 is a non-null subject language (Chini 2005; Torregrossa and Bongartz 2018). d) they overuse lexical DPs in their non-null subject L2 when their L1 is a non-null subject language (Leclercq and Lenart, 2013) but also when their L1 is a null subject language (Ryan 2015). A generalization that seems to emerge in this variegated picture is the following:

(2) a. Whenever overuse of overt pronouns by LBs is observed, a null subject language is at least involved.
b. Whenever overuse of lexical DPs by LBs is observed, a non-null subject language is at least involved.[15]

[15] As far as we know, the only exception to this generalization is reported in Torregrossa *et al.* 2017, with Greek-Albanian bilingual children overusing lexical DPs. Adult-like use and interpretation of the referential properties of pronouns, however, is a late acquired competence in monolingual and bilingual children (see Sorace *et al.* 2009; Papadopoulou *et al.* 2015 a. o.), so we believe that this finding does not represent a counterevidence to the generalization in (2), which is intended as pertaining to LBs.

As we have seen, each author has proposed an explanation for her data: an explanation which seems, however, to conflict, at least in part, with the data brought to by other studies.

Cross-linguistic influence, in particular the influence of English into Italian, could explain the data of Tsimpli *et al.* (2004), Sorace and Filiaci (2006) and Belletti, Bennati and Sorace (2007). Recall that English was the L2 of Tsimpli *et al.*'s speakers and the L1 of Sorace and Filiaci's and Belletti, Bennati and Sorace's speakers. The idea that cross-linguistic influence is the determining factor causing the overuse/over-acceptance of overt pronouns, however, does not explain why LBs with two null subject languages overuse/over-accept overt pronouns. Furthermore, cross-linguistic influence does not explain why overt pronouns are overused by LBs of a non-null subject language with a null subject L1. As Contemori and Dussias (2016) explicitly argue, if there were cross-linguistic influence effects, the data would reveal a high rate of null pronouns in the English productions of L1 Spanish LBs, contrary to fact. As highlighted since early studies concerning anaphora resolution in bilinguals, however, divergent patterns which involve proper syntactic violations (such as the use of null subjects in a non-null subject language) are not expected in near-native speakers (Tsimpli *et al.* 2004; Sorace and Filiaci 2006; Belletti, Bennati and Sorace 2007). Therefore, it is not surprising that we do not find null pronouns in the productions of this experimental group: these L2 speakers have acquired that there is no zero option in their L2 (i.e. they have correctly set the null subject parameter to the L2 value), and that, consequently, overt pronouns (or lexical DPs) must be used to pick up the current topic. What is surprising in Contemori and Dussias' data is that their LBs are, unexpectedly, under-explicit: they use less lexical DPs than English native speakers, particularly in the condition with two referents which differ for gender. This fact can be explained assuming that overt pronouns are used as a default form by these speakers, but the idea that overt pronouns are a default form for LBs does not explain the conspicuous amount of data showing the opposite pattern, i.e. a higher use of lexical DPs by LBs with respect to native speakers. Chini (2005) has found both overuse of overt pronouns and overuse of lexical DPs in the production of her L1 German-L2 Italian speakers, and she attributes this pattern to an influence of German into Italian. Cross-linguistic influence is instead excluded by Torregrossa and Bongartz (2018) as the factor determining an overuse of lexical DPs in the Italian productions of their German dominant bilinguals, since, as the authors argue, lexical DPs are attested both in German and in Italian. Their data are better explained, as they argue, through a processing account of bilinguals' difficulties with referring expressions. The question is, however, what kind of processing difficulties are bilinguals (LBs in particular) faced with. One possibility is that, as suggested by Sorace (2011, 2016 in particular), their difficulty is precisely in computing topicality.

In Di Domenico, Baroncini and Capotorti (2020) we analysed the subject anaphoric devices produced by three groups of speakers (Greek Natives, Italian Natives and L1 Greek-L2 Italian near-native LBs). Participants had to watch a short silent movie and then tell the story. We first counted the amount of *pros*, overt pronouns and lexical DPs produced by the three groups of speakers and we found no differences in the amount of lexical DPs produced. We also found that overt pronouns were significantly overused by LBs (both with respect to Greek natives and Italian natives), in line with much current literature. Finally, we found a nearly significant higher amount of *pros* in Greek natives with respect to the other two groups. We then analysed the occurrence of the different kinds of anaphoric devices in terms of topicality. The analysis revealed that the higher use of *pro* by Greek natives was properly significant in topic shift contexts: *pro* in Greek is more used in topic shift contexts than in Italian (see Torregrossa, Andreou and Bongartz 2020 for similar findings), and LBs do not transfer this property of their L1 into their L2. This means that they are able to compute topicality.

We also found that the overuse of overt pronouns by LBs was highly significant in topic continuity, again in line with much current literature. We further analysed the occurrence of overt pronouns in terms of number and kind of active referents, and we found that all groups of speakers use overt pronouns particularly when there are two active referents which differ in gender (and/or number), and LBs at a significantly higher rate with respect to the two groups of native speakers. Most of the overt pronouns used in topic continuity by LBs were in this specific context.

We interpreted these findings as follows: LBs have difficulties in computing topicality in contexts in which there is more than one active referent. As Arnold and Griffin (2007) have shown, the presence of an additional character decreases referents activation, and English (native) speakers tend to use more explicit devices in contexts with two referents. Going back to our LBs, in these contexts, decreased referents' activation makes the topic/non-topic distinction unclear, and so overt pronouns are the right device to pick up a referent whose status is unclear in terms of topicality.

Their over-explicitness, thus, is just outward. Furthermore, they do not have general difficulties in computing topicality, since they use *pro* in their L2 Italian as Italian native speakers do.

An explanation in terms of decreased activation, furthermore, naturally explains why overt pronouns may be used by LBs of two null subject languages in the place of *pro,* but not the reverse.

Our data do not support the view that overt pronouns are used as a default option by LBs, because they are used in the same contexts (2 referents with a gender and/or number mismatch) in which they are used by native speakers. Overt pronouns, furthermore, are not used to avoid lexical retrieval, since a comparable amount of lexical DPs is used by the three groups of speakers, as we have seen.

On the basis of Di Domenico, Baroncini and Capotorti (2020), we thus formulate the Decreased Activation Hypothesis:

(3) a. The Decreased Activation Hypothesis (DAH)
In the presence of two (or more) potential referents, their level of activation decreases.
In this condition, LBs treat each referent as non-prominent (non-topical).

In the next sections, we will explore the possibility that the DAH can be extended to account for all the data concerning LBs reviewed so far. Before doing so, however, we will try to find an explanation for the generalization in (2), which may suggest a difference in the way subject anaphoric devices are chosen/interpreted in null and non-null subject languages.

3. The choice and interpretation of subject anaphoric devices in null and non-null subject languages

Languages avail themselves of different devices to pick up a referent that has been previously introduced in the discourse. An interesting question is how speakers make their choice among the various possibilities made available by their language. A basic tenet shared by models stemming from different theoretical backgrounds, such as Grice's (1975) Maxim of Quantity, Chomsky's (1981) Avoid Pronoun principle, Cardinaletti and Starke's (1999)"Choice of a Pronoun" principle and Ariel's (1990, 2001) Accessibility Theory, states that a device should be as minimal as possible.

According to Ariel (1990, 2001) referring expressions (NP types, in her terms) are ranked on a universal scale of accessibility which proceeds from low (more explicit) to high (less explicit) accessibility markers.

Each referring expression encodes a specific degree of mental accessibility of its referent: the less accessible a referent is, the more conspicuous/explicit (in terms of informativity, rigidity and degree of attenuation), the referential marker used by the speaker will be. Lexical DPs, thus, are ranked higher than overt pronouns, which are in turn ranked higher than null pronouns. Wherever two forms are available in a language, they are assumed to occur in the same hierarchical order cross-linguistically. Referents' accessibility is in turn a function of factors such as salience, competition, and distance. [16]

Cardinaletti and Starke's (1999: 154) Semantic Asymmetry #1 similarly states that deficient (i.e. weak and clitic) pronouns must have an antecedent prominent in the discourse. [17]

Although there is perhaps still much to be done to establish the factors determining prominence (see also Cardinaletti and Starke 1999, fn. 11), there is general consensus on factors such as syntactic position (the subject is more prominent than the object) and topicality (the topic is more prominent than the non-topic).

It is well known that languages differ in the syntactic constraints that allow subject pronouns. The availability of null subjects (*pros*) is possible due to the positive setting of the null subject parameter (Taraldsen 1980; Rizzi 1982). Null subjects are not allowed in English and other non-null subject languages, with very limited exceptions such as ellipses and coordinated structures or some peculiar written and oral registers (Haegeman 2007; Weir 2009). An interesting question is whether this difference entails a different use of overt pronouns in null and non-null subject languages. If, however, overt pronouns in non-null subject languages can be weak or strong, and *pro* is just the null analogue of a weak pronoun (Cardinaletti and Starke 1999; Cardinaletti 2004a, 2004b, see also fn. 6 above) this difference is not predicted by the grammar.[18]

Still, languages may differently employ the possibilities offered by their grammar, and in what follows we will examine how subject anaphoric devices are used in null and non-null subject languages.

We will take into account mainly Italian (as representative of null subject languages) and English (as representative of non-null subject languages), assuming that, *modulo* some cross-linguistic micro-variation (for some examples see Section 2 above, as well as Di Domenico and Matteini 2021, a. o.) these considerations can be extended among the languages of each group.

A first issue we will explore concerns the division of labour between null and overt subject pronouns in Italian.

Seminal work by Calabrese (1986a, 1986b) has highlighted that, when there are two grammatically possible antecedents for a pronoun, null pronouns in Italian co-refer with/signal a topical antecedent (the "subject of primary predication", or Thema, in Calabrese's terms), which is in turn the "expected" antecedent. Overt pronouns ("stressed" in Calabrese's terms) instead co-refer with/signal an "unexpected", non-topical antecedent.

[16] A referent is less accessible if there are competing active referents or if it is more distant from the device that picks it up. As for salience: "the more salient the antecedent, the more highly accessible it is deemed." (Ariel 2001: 32).

[17] Strong and deficient pronouns, however, are not necessarily morphologically and phonologically distinct for Cardinaletti and Starke (1999).

[18] *pro* shares with its overt (weak) counterparts all syntactic and semantic properties but has different phonological properties (Cardinaletti 2004a: 132).

So, in (4) the null pronoun (*pro*) is interpreted as co-referent with *Carlo*, while the overt pronoun (*lui*) is interpreted as co-referent with *Sandro*:

(4) a. Dopo che Carlo$_i$ rimproverò Sandro$_j$ *pro*$_i$/ lui$_j$ cominciò a piangere
after that C. reproached S. *pro*/ he began to cry
'After Carlo reproached Antonio, he began to cry'
(*Some Properties of the Italian Pronominal System: An Analysis Based on the Notion of Thema as Subject of Predication,* Calabrese 1986b: 26)

This effect stems from the combined effect of the two principles in (5) and (6):

(5) a. Assign the feature [+stressed] to a pronominal X only when the occurrence of the referent of X is not expected
(*Pronomina*, Calabrese 1986a: 7)

(6) a. A pronominal in position of Thema is expected to have a referent of another Thema
(*Some Properties of the Italian Pronominal System: An Analysis Based on the Notion of Thema as Subject of Predication,* Calabrese 1986b: 31)

Experimental evidence provided by Carminati (2002) confirms Calabrese's analysis. Through a series of experiments testing the interpretation of overt and null pronouns in bi-clausal sentences analogous to (4), she shows that in Italian *pro* looks for an antecedent in a prominent syntactic position, i.e. the preverbal subject position (Spec, IP for Carminati 2002).[19, 20]

Another use of the overt pronoun in Italian, of a different nature, is when focalization is involved, as in (7) below:

(7) a. Lui ha fatto questo (non Carlo)
he has done this (not C.)
'He did this (not Carlo)'
(*Pronomina*, Calabrese 1986a: 9)

[19] See Cardinaletti (2004a) for the idea that *pro* and overt pronouns occupy two distinct subject positions. As noted by Cardinaletti (2004a: 149) a principle like (6) could not fit her analysis, *pro* occupying a subject position (spec, AgrSP) distinct from the position reserved to subjects of predication (spec, SubjP).

[20] It is worth noting that Carminati (2002: 34) explicitly assumes that the subject is the topic of the sentence, and this might well be the case in her experimental materials: in bi-clausal sentences like (4), without a context, we expect the subject (and not the object) to be interpreted as the topic. Calabrese's (1986a, 1986b) Thema is partially different from 'subject' as well as from 'topic'. He notes indeed that a post-verbal subject cannot be the antecedent of a pronoun (either null or overt), as shown in (i), while the benefactive dative argument of verbs of the *piacere* type in Italian are the preferred antecedents of *pro*, as shown in (ii):

(i) Quando è arrivato Carlo$_i$, *pro*$_i$/lui$_{*i}$ ha parlato
when is arrived Carlo pro/ he has spoken

(ii) Poiché a Carlo$_i$ è piaciuta Maria$_j$, *pro*$_{i/*j}$ vuole rimanere qui
since to Carlo is pleased Maria, pro wants remain here

A left-dislocated topic, instead, is not a possible antecedent of the null pronoun, as shown in (iii):

(iii) Poiché a Mario$_i$ Carla$_j$ gli ha dato un bacio, *pro*$_{*i/j}$ è felice
since to Mario Carla him has given a kiss, *pro* is happy

We will assume therefore that *pro* looks for an antecedent which is both a topic and a subject, i.e. an 'aboutness subject' in Rizzi's (2018) terms.

Since null pronouns cannot be focalized, when focalization is involved null pronouns are excluded, so that there is no possible division of labour between null and overt pronouns under focalization.

Leaving aside the case of focalization, thus, there seems to be a division of labour between null and overt pronouns in Italian, with null pronouns picking up the topical subject, and overt pronouns the non-topic, non-subject antecedent (Calabrese 1986a, 1986b; Carminati 2002).

These biases however appear to be less clear in some peculiar syntactic environments. One of these environments is discussed in Rizzi (2018), and concerns the interpretation of *pro* in bi-clausal main/complement clauses like (8):

(8) a. Francesca ha fatto notare a Maria che *pro* era molto stanca
Francesca has made notice to Maria that *pro* was very tired
'Francesca made Maria realize that (she) was very tired'

Here, the preference for the subject/topic antecedent *Francesca* (the aboutness subject in Rizzi's terms) is by far less clear, with *pro* being able to co-refer with *Maria* as well. Rizzi (2018) assumes that c-command weakens what he calls the "Calabrese effect", proposing an extra principle like (9):

(9) a. A subject pronoun is expected to have the referent of a c-commanding DP
(*Subjects, Topics and the Interpretation of Pro*, Rizzi 2018: 516)

As c-command makes every DP a potential antecedent for a pronoun, the Calabrese effect is visible when there is no c-command, i.e. with pre-posed adjunct adverbial clauses (as in (4) above) and in discourse sequences.

This might also explain the weaker subject bias of the null pronoun detected by some studies (Tsimpli *et al.* 2004; Sorace and Filiaci 2006; Belletti, Bennati and Sorace 2007 a. o.) in bi-clausal constructions such as (10.a), with respect to (10.b):

(10) a. L' anziana signora saluta la ragazza quando *pro* attraversa la strada
the old lady greets the girl when *pro* crosses the street
b. Quando *pro* attraversa la strada, l' anziana signora saluta la ragazza
when *pro* crosses the street the old lady greets the girl

In sentences like (10.a), Belletti, Bennati and Sorace (2007) found a preference of the 40% only for *l'anziana signora* as the antecedent of *pro*. In (10.b) the same preference reaches 90%. The weak preference for the subject antecedent in (10.a) can be explained assuming that in this configuration not only the subject, but also the object antecedent (Pesetsky 1995) c-commands *pro*, while in (10.b) they do not, and hence the "Calabrese effect" is visible.

Belletti, Bennati and Sorace (2007) also note that the interpretation of the overt pronouns differs in (11.a) and (11.b):

(11) a. L' anziana signora saluta la ragazza quando lei attraversa la strada
the old lady greets the girl when she crosses the street
b. Quando lei attraversa la strada, l' anziana signora saluta la ragazza
when she crosses the street the old lady greets the girl

While the interpretation of the overt pronoun as co-referent with the object antecedent (*la*

ragazza) is highly preferred in (11.a), i.e. 85%, it only reaches 23% in (11.b). In the latter case, the overt pronoun is interpreted as co-referent with an external referent in the 57% of the cases.

Here, we propose, the preference for an external referent possibly stems from the fact that constructions like (11.a) are frequently used in (semi-spontaneous as well as elicited) production with a specific purpose/meaning. In two production studies (Di Domenico, Baroncini and Capotorti 2020; Contemori and Di Domenico 2021), we found examples like (12) and (13):

(12) a. mentre lui è sceso dalla scala [.....]
while he is climbed-down from-the ladder
b. passa un, *pro* presumo, allevatore
passes a *pro* guess1S farmer
'While he is down the ladder a farmer, I guess, passes by.'

(13) a. Minnie è andata con Paperina al Parco Nazionale d'Abruzzo.
Minnie went with Daisy to Abruzzo National Park
b. Mentre lei scatta delle foto, Paperina passeggia.
While she takes pictures, Daisy walks around

In (12), the current topic, talked about in several previous sentences, is a peasant who is picking pears. The overt pronoun in (12.a) co-refers with this antecedent, but establishes a contrast with the referent introduced immediately after. Similarly, in (13.b) the overt pronoun co-refers with the subject antecedent of (13.a), but establishing a contrast with the other character. It is possible that in sentences like (11.b) the external referent bias of the overt pronoun instantiates the interpretive characteristics that overt pronouns have in (12.a) and (13.b).

Another use of the overt pronoun is described by Cardinaletti (2004a, 2004b) and Frascarelli (2007), where an overt (weak) pronoun can be used to restate the current topic (which is perhaps too distant) and can have a non-human antecedent, as *loro* in (14) below, taken from the novel *Novecento* by A. Baricco (adapted from Cardinaletti 2004b: 141):

(14) a. (i quadri) *pro* stanno lì attaccati al chiodo, nessuno gli fa niente, ma loro a un certo punto, *fran*, cadono giù come sassi
(the paintings) They stay there, hanged to the nail, nobody does them anything, but they, suddenly, *fran,* fall down as stones

To sum up, in Italian a null pronoun co-refers with an aboutness subject antecedent, provided c-command does not interfere. An overt pronoun signals instead a topic shift, or a contrast as in (7), (12) and (13). In some cases, such as (14), it is used to restate the current topic, the latter an emerging peculiarity quite rarely attested and subject to individual variation (Cardinaletti 2004a, 2004b; Frascarelli 2007). As noted by Calabrese (1986b) the expectation is that one will continue to talk about the current subject of predication: hence, the current aboutness subject is the most salient (i.e. accessible) antecedent (see also Crawley 1986; Carminati 2002; Colonna, Schimke and Hemforth 2012 a. o.). The division of labour between null and overt pronouns in Italian thus appears congruent with Ariel's (1990, 2001) Accessibility Theory as well as with Cardinaletti and Starke's (1999) Semantic Asymmetry.

Di Domenico, Baroncini and Capotorti (2020) also show that overt pronouns are a marked choice in the production of Italian Natives, since they instantiate the 6.26% only of the anaphoric devices attested in the corpus. The analysis in terms of topicality and number

and kind of active referents confirms that overt pronouns are by far more used in topic shift contexts, and virtually absent in topic continuity. [21] In topic shift contexts, overt pronouns are significantly more used in the condition in which there are two active referents (i.e. two referents explicitly mentioned in the considered clause and/or in the clause preceding it) with a gender (or number) mismatch (14/24), and never in the condition with 1 referent. The 1 referent condition is instead the condition in which lexical DPs are mostly employed to realize a topic shift (in the 34 cases of topic shifts with 1 referent, 27 lexical DPs were employed). Since in this situation the antecedent of the topic-shifting anaphoric device is more distant, we may indeed assume that there is a ranking, with lexical DPs used to retrieve the less accessible antecedents and overt pronouns the more accessible ones. Similar findings concerning overt pronouns and lexical DPs characterize the productions of Greek Natives.[22] Finally, lexical DPs may also be used in topic continuity in Italian and Greek.

Let us now turn to English. Some authors (see e.g. Azar, Özyürek and Backus 2019 and the references quoted there) claim that in English, and in non-null subject languages in general, the topic/non topic distinction is not deemed as very relevant, and an overt pronoun can pick up either a topical as well as a non- topical antecedent. As we have seen, along the lines of Cardinaletti and Starke (1999) and Cardinaletti (2004a, 2004b) instead, overt pronouns in English can be weak – and so they pick up a topical antecedent – or strong, signalling a topic shift.[23]

In (15) we report an example taken from Michaelis and Francis (2007) which illustrates these differences, with topic shifting pronouns in bold and topic continuing ones in italics:

(15) a. My sister has a, *she* just had a baby.
b. **He**'s about five months old,
c. and **she** was worrying about going back to work and
d. what *she* was going to do with him.

Various experimental and corpus studies seem to suggest, however, that the preferred interpretation of overt pronouns in English is towards the subject/topic antecedent, while in production overt pronouns are used in topic continuity, particularly when there is a single active referent (Arnold and Griffin 2007). Using the weak/strong distinction, this would be equivalent to say that in English weak pronouns are more widespread than strong pronouns.

Arnold (2010) proposed the term "first-mentioned bias" to account for this preference.

Arnold *et al.* (2000) for instance, conducted an eye-tracking study where participants viewed a picture while listening to a four-sentences story describing it. The first sentence introduced the two characters (either of the same gender or of different genders), while the third sentence contained a pronoun which referred either to the first – or the second – mentioned character, for a total of four conditions. They found that participants did not immediately converge on

[21] Six of the seven overt pronouns in the corpus which are not in topic shift contexts encode either focalization or the kind of contrast involved in (11) and (12) above, or are an instance of the weak, emphatic pronoun described by Cardinaletti (2004a, 2004b) and Frascarelli (2007).

[22] Lozano (2016) found similar data concerning the distribution of overt pronouns in Spanish Natives' productions.

[23] The correlation of the weak/strong distinction with prosody is still unclear. Earlier studies (e.g. Sheldon 1974) suggest that when a pronoun is 'accented' it is meant to refer to the non-subject antecedent. Similarly, Calabrese (1986) proposes that overt pronouns in Italian are equivalent to English stressed pronouns. According to Gargiulo, Tronnier and Bernardini (2019) however, prosodic prominence of the pronoun signals its non-preferred interpretation, both in Swedish and in Italian.

an interpretation of the pronoun referent when the pronoun referred to the second-mentioned character in the same gender condition. In this condition, participants did not look at the target more than the competitor, as they did in the other conditions, where either a gender or an accessibility cue was present. Corpus studies, furthermore, report a relatively high frequency of lexical DP subjects (particularly in written registers) with a "switch topic" function (Michaelis and Francis 2007 a.o.).

This pattern seems to be confirmed also in non-null subject languages other than English.

Gargiulo, Tronnier and Bernardini (2019) using sentences analogous to (1.d) have shown that in Swedish overt pronouns are significantly more interpreted as co-referent with a subject/topic antecedent (66% *vs.* 44%).

Hendriks, Koster and Hoeks (2014) examined Dutch speakers in comprehension and production. Restricting our attention to their young adults' group, overt pronouns were preferably used for topic maintenance in contexts with only one active referent. In the condition with two referents of the same gender, participants preferred a full DP to refer to the topical antecedent as well as to the non-topic. In comprehension, overt pronouns were interpreted as referring to the topic, i.e. to the subject of the previous clause. [24]

To sum up, while overt pronouns in Italian (and also in Greek and possibly Spanish) are typically used in co-reference with a non-topical antecedent, lexical DPs are preferred in English (and Dutch) in the same contexts, despite the fact that overt pronouns may also be strong in these languages.[25] Overt pronouns are more widespread as topic continuing devices, i.e. as weak pronouns, in English, Dutch and Swedish.

4. Two ways of being "seemingly over-explicit"

In 2.3, following Di Domenico, Baroncini and Capotorti (2020), we have proposed the Decreased Activation Hypothesis (here repeated for convenience) to explain the overuse of overt pronouns by LBs of two null subject languages, Greek and Italian:

(3) a. The Decreased Activation Hypothesis (DAH)
In the presence of two (or more) potential referents, their level of activation decreases. In this condition, LBs treat each referent as non-prominent (non-topical).

Assuming that when more than one referent is active the topic/non topic distinction is unclear (i.e. all referents are treated as if they were non-topical), LBs coherently use an overt pronoun to pick up a previously introduced referent (particularly when there is a gender/number mismatch, as native speakers of Italian, Spanish and Greek do).

Can the DAH be extended to explain the data reported by the studies reviewed in Section 2?

[24] Hendricks, Koster and Hoeks (2014) found that young adult speakers of Dutch used more lexical DPs with two active referents when compared to children and elderly adults. It is worth noting that their materials contained two referents of the same gender and number, and that use of a pronoun would be ambiguous in this context in Dutch or English. Their interpretation of this result is different from Arnold and Griffin's: they assume that young adults are able to take into account the listener's perspective, avoiding the ambiguous pronoun. In Italian, an overt pronoun would not be ambiguous in this context as example (2) shows.

[25] As an anonymous reviewer notes, it may be the case that the overt pronoun in (at least canonical) null subject languages reinforces the features of the inflectional verbal affix (adding further specifications such as gender), while in non-null subject languages, given that inflectional agreement is limited or absent, the overt DP provides full specification.

The generalization in (2), here repeated for convenience, reveals that whenever overuse of overt pronouns is observed, a null subject language is at least involved, and whenever overuse of lexical DPs is observed, a non-null subject language is at least involved:

(2) a. Whenever overuse of overt pronouns by LBs is observed, a null subject language is at least involved.
b. Whenever overuse of lexical DPs by LBs is observed, a non-null subject language is at least involved.

As we have seen in Section 3, furthermore, overt pronouns are typically used/ interpreted in co-reference with a non-topical antecedent in null subject languages, while in non-null subject languages overt pronouns typically pick up a topical antecedent, and lexical DPs are preferred when a non-topical antecedent is addressed.

Let us assume that bilingual (or multilingual) competence implies the possibility of having at disposal a wider range of options with respect to monolingual competence (Belletti, Bennati and Sorace 2007; Di Domenico 2015; Caloi, Belletti and Poletto 2018) and that bilingual speakers (particularly if they are very advanced in both languages) are free to choose among these options, provided that their choices do not violate the grammatical constraints of the target language.

Assuming that decreased activation is the specific problem shared by all LBs, a speaker of a null and a non-null subject language can cope with this problem either overusing overt pronouns (the device specialized to pick up a non-topical antecedent in null subject languages) or overusing lexical DPs (the device preferred in non-null subject languages to pick up a non-topical antecedent). The German L1 – Italian L2 speakers investigated by Chini (2005) seem to opt for both solutions, in that they overuse overt pronouns and also lexical DPs. The L1 Italian – L2 English speakers studied by Tsimpli *et al.* (2004), the L1 English- L2 Italian speakers of Belletti, Bennati and Sorace (2007) and the L1 Spanish-L2 English speakers of Contemori and Dussias (2016) opt instead for the overt pronouns solutions, while the L1 Chinese – L2 English speakers studied by Ryan (2015) and the German-Italian bilinguals of Torregrossa and Bongartz (2018) opt for the lexical DPs solution. Torregrossa and Bongartz (2018) state that cross-linguistic influence cannot be the factor causing overuse of lexical DPs, since lexical DPs are attested both in Italian and in German: extending to German the considerations outlined in Section 3 for English, Dutch and Swedish, in German lexical DPs might be preferred to co-refer with a non-topical antecedent, so cross-linguistic influence might be implied in this result, though intended as a less automatic process than commonly assumed.

The lexical DPs solution is not adopted by speakers of two null subject languages (Margaza and Bel 2006; Di Domenico and Baroncini 2019) since overt pronouns are typically employed in null subject languages to pick up a non-topical antecedent. Similarly, the overt pronoun solution is not adopted by speakers of two non-null subject languages (Leclercq and Lenart 2013) since overt pronouns in non-null subject languages are not the preferred device to pick up a non-topical antecedent.

Crucially, however, the DAH predicts that we should not observe an over-explicit form when only one referent is active, as in this example from Italian, reported by Sorace (2016: 4):

(16) a. Perché Maria è arrivata così tardi?
why Maria is arrived so late
'Why did Maria arrive so late'
b. Perché lei si era addormentata.

because she REFL*si* was asleep
'Because she fell asleep'

In this condition, one cannot reasonably maintain that *Maria* is not treated as a topic, and there is no reason why its activation should decrease. Unfortunately, it is impossible to check this feature in all the studies reviewed in Section 2, since an analysis in terms of active referents was not performed by the authors.[26]

Lozano (2016) reports that cases like (16) are very rare in the L2 Spanish productions of his L1 English speakers, which produced redundant overt pronouns mostly in contexts with two or more referents. In Di Domenico, Baroncini and Capotorti (2020), they were quite rare as well: 4 cases, three of which produced by the same speaker (which was a very advanced speaker of English, her L3).

Contemori and Dussias (2016), as we have seen, report however an overuse of overt pronouns also in the 1 referent condition, although less consistent than in the condition with two referents differing for gender.

These cases, which are not explained by the DAH, are possibly an instance of a more "direct" process of cross-linguistic influence from English (the L1 in Lozano's speakers, the L2 in Contemori and Dussias', the L3 in Di Domenico, Baroncini and Capotorti's speaker) in which the properties of an element in Language A are transferred to the 'equivalent' element in Language B (see Di Domenico 2015, 2020 for a description of this process): the subject bias of the English (weak) overt pronoun is transferred to the Italian (strong) overt pronoun.

5. Conclusions

In this work we have proposed the Decreased Activation Hypothesis as a unifying explanation for a set of *prima facie* conflicting data, which have been explained in partially conflicting ways, concerning the way in which bilinguals (in particular late bilinguals) choose and interpret anaphoric devices. We have proposed that the specific processing problem that LBs experience concerns referents' activation, which decreases in contexts with more than one referent, so that LBs treat all referents as non-topical, using a seemingly over-explicit device. This seemingly over-explicit device, as reported by the relevant literature, can be an overt pronoun or a lexical DP. The first step to explain this different choice has been the observation that null subject languages use preferably an overt pronoun to pick up a non-topical antecedent, while non-null subject languages use preferably a lexical DP to pick up a non-topical antecedent (despite the fact that the strong pronoun option is available also in non-null subject languages). Then, considering that whenever overuse of overt pronouns is reported at least one of the languages of the bilinguals is a null subject language, and whenever overuse of lexical DPs is reported at least one of the languages of the bilinguals is a non-null subject language, we have proposed that bilinguals of a null and a non-null subject language can choose one of the two options offered by their multilingual competence. With respect to other explanations proposed in the previous literature, we include in our proposal both processing difficulties (as in Sorace 2011, 2016) and cross-linguistic influence (as in earlier proposals as Tsimpli *et al*. 2004; Belletti, Bennati and Sorace 2007), although a kind of cross-linguistic influence interpreted as a choice

[26] The materials in the Picture Verification task used by Tsimpli *et al.* (2004) and Belletti, Bennati and Sorace (2007), however, always contained two referents. Lozano (2018) had always more than one referent in his experimental sentences. Ryan (2015) reports over-explicitness particularly attested in contexts with less accessible referents.

among different possibilities offered by multilingual competence. We do not assume that overt pronouns are a default choice for LBs, otherwise overuse of lexical DPs would remain puzzling. Last but not least, we assume that LBs often claimed over-explicitness is just outward.

As a final remark we note that in order to strengthen our proposal it would be necessary to compare speakers of different combinations of languages with similar experimental materials, which in particular distinguish the conditions with one referent from the conditions with two or more referents, as well as the conditions of topic continuity from the conditions of topic shift. Perhaps this problem (comparability of experimental results) is more general in the field of anaphora resolution in bilinguals, and beyond: we hope to overcome it in future research.

References

Abutalebi, Jubin. 2008. "Neural Aspects of Second Language Representation and Language Control." *Acta Psichologica* 128 (3): 466-478.

Ariel, Mira. 1990. *Accessing Noun-Phrase Antecedents.* London: Routledge.

Ariel, Mira. 2001. "Accessibility Theory: An Overview." In *Text Representation: Linguistic and Psycholinguistic Aspects,* ed. by Ted J.M. Sanders, Joost Schilperoord, and Wilbert Spooren, 29-87. Amsterdam: John Benjamins Publishing Company.

Arnold, Jennifer E., Janet G. Eisenband, Sarah Brown-Schmidt, and John C. Trueswell. 2000. "The Rapid Use of Gender Information: Evidence of the Time Course of Pronoun Resolution from Eyetracking." *Cognition* 76 (1): 13-26.

Arnold, Jennifer E., and Zenzi M. Griffin. 2007. "The Effect of Additional Characters on Choice of Referring Expression: Everyone Counts." *Journal of Memory and Language* 56 (4): 521-536.

Arnold, Jennifer. 2010. "How Speakers Refer: The Role of Accessibility." *Language and Linguistics Compass* 4 (4): 187-203.

Azar, Zeynep, Aslı Özyürek, and Ad Backus. 2019. "Turkish-Dutch Bilinguals Maintain Language Specific Reference Tracking Strategies in Elicited Narratives." *International Journal of Bilingualism* 24 (2): 376-409.

Belletti, Adriana, Elisa Bennati, and Antonella Sorace. 2007. "Theoretical and Developmental Issues in the Syntax of Subjects: Evidence from Near-native Italian." *Natural Language and Linguistic Theory* 25 (4): 657-689. doi: 10.1007/s11049-007-9026-9.

Bialystock, Ellen. 2009. "Bilingualism: The Good, the Bad, and the Indifferent." *Bilingualism: Language and Cognition* 12 (1): 3-11.

Bini, Milena. 1993. "La adquisicíon del italiano: mas allá de las propiedades sintácticas del parámetro pro-drop." In *La Linguistica y el Analisis de los Sistemas no Natives*, ed. by Juana Liceras, 126-140. Ottawa: Doverhouse Editions Canada.

Calabrese, Andrea. 1986a. "Pronomina." In *Papers in Theoretical Linguistics*, ed. by Fukui Naoki, Tova R. Rapoport, and Elizabeth Sagey, 1-46. Cambridge, MA: MIT Press.

Calabrese, Andrea. 1986b. "Some Properties of the Italian Pronominal System: An Analysis Based on the Notion of Thema as Subject of Predication." In *Tema/Rema in italiano*, ed. by Harro Stammerjohan, 25-36. Tübingen: Gunter Narr Verlag.

Caloi, Irene, Adriana Belletti, and Cecilia Poletto. 2018. "Multilingual Competence Influences Answering Strategies in Italian-German Speakers." *Frontiers in Psychology*. doi: 10.3389/fspsg.2018.01971.

Cardinaletti, Anna, and Michael Starke. 1999. "The Typology of Structural Deficiency. A Case Study of the Three Classes of Pronouns." In *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 145-233. Berlin-New York, NY: Mouton de Gruyter.

Cardinaletti, Anna. 2004a. "Toward a Cartography of Subject Positions." In *The Structure of CP and IP. The Cartography of Syntactic Structures,* vol. II*,* ed. by Luigi Rizzi, 115-165. Oxford: Oxford UP.

Cardinaletti, Anna. 2004b. "La traduzione dei pronomi: interferenza sintattica e cambiamento linguistico." In *Lingua, mediazione linguistica e interferenza,* ed. by Giuliana Garzone, and Anna Cardinaletti, 129-166. Milano: Franco Angeli.

Carminati, Maria N. 2002. *The Processing of Italian Subject Pronouns*. PhD Dissertation, University of Massachusetts Amherst.
Chini, Marina. 2005. "Reference to Person in Learner Discourse." In *The Structure of Learner Varieties,* ed. by Henrietta Hendriks, 65-110. Berlin: Walter de Gruyter.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Clasen, Harald, and Claudia Felser. 2006. "Grammatical Processing in Language Learners." *Applied Psycholinguistics* 27 (1): 3-42. doi: 10.1017/S142716406060024.
Colonna, Saveria, Sarah Schimke, and Barbara Hemforth. 2012. "Information Structure Effects on Anaphora Resolution in German and French: A Crosslinguistic Study of Pronoun Resolution." *Linguistics* 50 (5): 991-1013.
Contemori, Carla, and Paola E. Dussias. 2016. "Referential Choice in a Second Language: Evidence for a Listener-oriented Approach." *Language, Cognition and Neuroscience* 31 (10): 1257-1272.
Contemori, Carla, and Elisa Di Domenico. 2021. "Microvariation in the Division of Labor Between Null and Overt Subject Pronouns: The Case of Italian and Spanish." *Applied Psycholinguistics 42* (4): 997-1028.
Crawley, Rosalind. 1986. "Some Factors Influencing the Comprehension of Pronouns in Text." In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* (Amherst, MA, 15-17 August 1986), ed. by Charles Clifton, 613-620. Hillsdale, NJ: Lawrence Erlbaum Associates.
Di Domenico, Elisa. 2015. "Transferring Strategies and the Nature of Transfer." In *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 237-255. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company.
Di Domenico, Elisa, and Ioli Baroncini. 2019. "Age of Onset and Dominance in the Choice of Subject Anaphoric Devices: Comparing Natives and Near-natives of Two Null-Subject Languages."*Frontiers in Psychology.* doi: 10.3389/fpsyg.2018.02729.
Di Domenico, Elisa. 2020. "To the Root of (Some) Interface Properties." In *Language Acquisition, Processing and Bilingualism. Selected Papers from the Romance Turn VII*, ed. by Anna Cardinaletti, Chiara Branchini, Giuliana Giusti, and Francesca Volpato, 160-182. Newcastle upon Tyne: Cambridge Scholars Publishing.
Di Domenico, Elisa, Ioli Baroncini, and Andrea Capotorti. 2020. "Null and Overt Subject Pronouns in Topic Continuity and Topic Shift: An investigation of the Narrative Productions of Italian Natives, Greek Natives and Near-native Second Language Speakers of Italian with Greek as a First Language." *Glossa: A Journal of General Linguistics* 5 (1): 1-35. doi: 10.5334/gjgl.1009.
Di Domenico, Elisa, and Simona Matteini. 2021. "Microvariation in Subject Realization and Interpretation: an Introduction." *Glossa. A Journal of General Linguistics* 6 (1): 1-11. doi: 10.5334/gjgl.1610.
Frascarelli, Mara. 2007. "Subjects, Topics and the Interpretation of Referential *Pro*. An Interface Approach to the Linking of (Null) Pronouns." *Natural Language and Linguistic Theory* 25 (4): 691-734.
Gargiulo, Chiara, Mechtild Tronnier, and Petra Bernardini. 2019. "The Role of Prosody in Overt Pronoun Resolution in a Null Subject Language and in a Non-null Subject Language: A Production Study." *Glossa: A Journal of General Linguistics* 4 (1). doi: 10.5334/gjgl.973.
Grice, Paul H. 1975. "Logic and Conversation." In *Syntax and Semantics vol. 3, Speech Acts*, ed. by Peter Cole, and Jerry L. Morgan, 41-58. New York, NY: Academic Press.
Haegeman, Liliane. 2007. "Subject Omission in Present-day Written English: On the Theoretical Relevance of Peripheral Data." *Rivista di Grammatica Generativa* 32: 91-124.
Hendriks, Petra, Charlotte Koster, and John C. J. Hoeks. 2014. "Referential Choice Across the Lifespan: Why Children and Elderly Adults Produce Ambiguous Pronouns." *Language, Cognition and Neuroscience* 29 (4): 391-407.
Judy, Tiffany. 2015. "Knowledge and Processing of Subject-related Discourse Properties in L2 Near-native Speakers of Spanish, L1 Farsi." In *The Acquisition of Spanish in Understudied Language Pairings*, ed. by Tiffany Judy, and Silvia Perpiñán, 169-199. Amsterdam: John Benjamins Publishing Company.
Kroll, Judith, and Ellen Bialystok. 2013. "Understanding the Consequences of Bilingualism for Language Processing and Cognition." *Journal of Cognitive Psychology* 25 (5): 497-514.

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

Leclercq, Pascale, and Ewa Lenart. 2013. "Discourse Cohesion and Accessibility of Referents in Oral Narratives: A comparison of L1 and L2 Acquisition of French and English." *Discours.* doi: 10.4000/discours.8801.

Lozano, Cristóbal. 2006. "The Development of the Syntax-information Structure Interface: Greek Learners of Spanish." In *The Acquisition of Syntax in Romance Languages,* ed. by Vincent Torrens, and Linda Escobar, 371-399. Amsterdam: John Benjamins Publishing Company.

Lozano, Cristóbal. 2016. "Pragmatic Principles in Anaphora Resolution at the Syntax-discourse Interface: Advanced English Learners of Spanish in the CEDEL2 Corpus." In *Spanish Learner Corpus Research: Current Trends and Future Perspectives,* ed. by Margarita Alonso-Ramos, 235-265. Amsterdam: John Benjamins Publishing Company.

Lozano, Cristóbal. 2018. "The Development of Anaphora Resolution at the Syntax-discourse Interface: Pronominal Subjects in Greek Learners of Spanish." *Journal of Psycholinguistic Research* 47 (2): 411-430.

Margaza, Panagiota, and Aurora Bel. 2006. "Null Subjects at the Syntax-pragmatics Interface: Evidence from Spanish Interlanguage of Greek Speakers." In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, (Banff, Canada, 27-30 April 2006), ed. by Mary Grantham O'Brien, Christine Shea, and John Archibald, 88-97. Somerville, MA: Cascadilla Press.

Michaelis, Laura A., and Hartwell S. Francis. 2007. "Lexical Subjects and The Conflation Strategy." In *The Grammar-Pragmatics Interface: Essays in Honor of Jeannette K. Gundel*, ed. by Nancy Hedberg, and Ron Zacharski, 19-48. Amsterdam: John Benjamins Publishing Company. doi: 10.1075/pbns.155.04mic.

Papadopoulou, Despina, Eleni Peristeri, Evagelia Plemenou, Theodoros Marinis, and Ianthi M. Tsimpli. 2015. "Pronoun Ambiguity Resolution in Greek: Evidence from Monolingual Adults and Children." *Lingua* 155: 98-120.

Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.

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

Rizzi, Luigi. 2005. "On Some Properties of Subjects and Topics." In *Proceedings of the 30th Incontro di Grammatica Generativa,* (Venice, Italy, 26-28 February 2004), ed. by Laura Brugè, Giuliana Giusti, Nicola Munaro, Walter Schweikert, and Giuseppina Turano, 203-224. Venezia: Cafoscarina.

Rizzi, Luigi. 2018. "Subjects, Topics and the Interpretation of *Pro*." In *From Sounds to Structures: Beyond the Veil of Maya*, ed. by Roberto Petrosino, Pietro Cerrone, and Harry van der Hulst, 510-529. Berlin: De Gruyter Mouton.

Ryan, Jonathon. 2015. "Over-explicit Referent Tracking in L2 English: Strategy, Avoidance or Mith?" *Language Learning* 65 (4): 824-859. doi:10.1111/lang.12139.

Sheldon, Amy. 1974. "The Role of Parallel Function in the Acquisition of Relative Clauses in English." *Journal of Verbal Learning and Verbal Behavior* 13 (3): 272-281.

Sorace, Antonella, and Francesca Filiaci. 2006. "Anaphora Resolution in Near-native Speakers of Italian." *Second Language Research* 22 (3): 339-368.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci, and Michela Baldo. 2009. "Discourse Conditions on Subject Pronoun Realization: Testing the Linguistic Intuitions of Older Bilingual Children." *Lingua* 119 (3): 460-477.

Sorace, Antonella. 2011. "Pinning Down the Concept of 'Interface' in Bilingualism." *Linguistic Approaches to Bilingualism* 1 (1): 1-33. doi: 10.1075/lab.1.1.01sor.

Sorace, Antonella. 2016. "Referring Expressions and Executive Functions in Bilingualism." *Linguistic Approaches to Bilingualism* 6 (5): 669-684. doi: 10.1075/lab.15055.sor.

Taraldsen, Tarald. 1980. *On the Nominative Island Constraint, Vacuous Application and the That-trace Filter*. Bloomington, IN: Indiana University Linguistic Club.

Toole, Janine. 1996. "The Effect of Genre on Referential Choice." In *Reference and Referent Accessibility,*

ed. by Jeanette K. Gundel, and Thorstein Fretheim, 263-290. Amsterdam: John Benjamins Publishing Company.

Torregrossa, Jacopo, Maria Andreou, Christiane Bongartz, and Ianthi M. Tsimpli. 2017. "Pinning Down the Role of type of Bilingualism in the Development of Referential Strategies." Paper presented at the Conference on *Generative Linguistics in the Old World (GLOW 40)* (Leiden, Netherlands, 14-17 March 2017. <https://glow2017.files.wordpress.com/2017/02/torregrossa-et-al-_alternate1.pdf> (06/2021).

Torregrossa, Jacopo, and Christiane Bongartz. 2018. "Teasing Apart the Effects of Dominance, Transfer and Processing in Reference Production by German-Italian Bilingual Adolescents." *Languages* 3 (3): 1-25. doi:10.3390/languages3030036.

Torregrossa, Jacopo, Maria Andreou, and Christiane Bongartz. 2020. "Variation in the Use and Interpretation of Null Subjects: A View from Greek and Italian." *Glossa: A Journal of General Linguistics* 5 (1).

Tsimpli, Ianthi M., Antonella Sorace, Caroline Heycock, and Francesca Filiaci. 2004. "First Language Attrition and Syntactic Subjects: A Study of Greek and Italian Near native Speakers of English." *International Journal of Bilingualism* 8 (3): 257-277.

Tsimpli, Ianthi M., and Antonella Sorace. 2006. "Differentiating Interfaces: L2 Performance in Syntax-Semantics and Syntax-Discourse Phenomena." In *Proceedings of the 30th Annual Boston University Conference on Language Development (BUCLD 30)*, ed. by David Bamman, Tatiana Magnitskaia, and Colleen Zaller, 653-664. Somerville, MA: Cascadilla Press.

Weir, Andrew. 2009. *Subject Pronoun Drop in Informal English*. PhD Dissertation, University of Edimburgh. Richard M. Hogg Prize winning essay. <http://people.umass.edu/aweir/> (06/2021).

White, Lydia, and Fred Genesee. 1996. "How Native is Near-native? The Issue of Ultimate Attainment in Adult Second Language Acquisition." *Second Language Research* 12 (3): 233-265.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: L. Franco (2021) The morphosyntactic interaction of kinship terms with evaluative morphemes in Italian. *Qulso* 7: pp. 75-88. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12003

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

The morphosyntactic interaction of kinship terms with evaluative morphemes in Italian

Ludovico Franco
Università degli Studi di Firenze (<ludovico.franco@unifi.it>)

Abstract:

In this brief paper, we address the interaction between evaluative morphology and kinship terms introduced by possessives in Italian, showing that the application of evaluative affixes influences the syntactic context in which kinship terms can be employed: they cannot be introduced by a bare determiner when evaluative morphemes attach to the lexical root. We argue that this (trivial) empirical observation has some clear consequences from a theoretical viewpoint: the fact that derivational morphemes, such as evaluatives, alter the syntactic environment in which a noun is couched supports the theory of grammar advanced in Manzini and Savoia (2007, 2011), who assume that Merge takes morphemes as its input and single morphemes are fully visible to the syntactic computation.

Keywords: *Determiners, Evaluative Kinship Terms, Morphology, Possessives, Proper Names*

1. Introduction: the empirical facts and some theoretical background

In Italian, there is an interesting asymmetry concerning the distribution of determiners and possessives when a kinship term is selected from the lexicon. With canonical kinship terms we may usually find sentences in which the determiner is (optionally) absent, as illustrated in (1a).[1] This is clear property of the set of kinship terms, given that (most) nominal items (including both common nouns and proper names) are obligatorily introduced by a D item in a corresponding syntactic position, as illustrated in (1b). The application of evaluative morphology on kinship terms (highlighted in bold below) yields interesting outcomes, given that the D item is now required, as illustrated in the examples in (1c).

[1] For this phenomenon, there is a full range of microvariation in the Dialects of Italy, which will not be considered in the present paper. The interested reader may refer to Cardinaletti and Giusti (2018), Baldi and Savoia (2019), Masi (2020), among others, for relevant data and analyses.

(1) a. (La) mia sorella/ (la) mia zia/ (la) mia nipote è andata al mare
'My sister/aunt/nephew(grandson) went to the sea'
b. *(Il) mio amico/*(il) mio Gianni/*(il) mio cane è andato al mare
'my friend/my Gianni/my dog went to the sea'
c. *(La) mia sorell**in**a/*(la) mia zi**ett**a/ *(la) mia nipot**in**a è andata al mare
'My younger sister/auntie /little nephew(grandson) went to the sea'

This is a trivial observation and is far from being novel. Longobardi (1994, 1996), Giusti (2015), Cardinaletti and Giusti (2018), among others, already have noticed the asymmetry between (1a) and (1c). Still, as far as I acknowledge, in the theoretical literature there are no attempts to formally explain the obligatory presence of D items in DP headed by kinship terms, once evaluative morphemes are attached to these nominal items. The aim of this brief paper is precisely to fill this gap providing a formal explanation for these facts.

Longobardi (1994 and subsequent literature) has shown that some classes of singular nouns in Italian can move leftward, triggering a N-to-D chain, raising from their base position to the one usually occupied by determiners, crossing over potentially intervening lexical items (e.g. possessive items). This is true, for instance, for many proper names and the item *casa* (home). In (2)-(3), it is possible to see that the determiner is absent only when the nominal item moves leftward.

(2) a. L'antica Roma fu la città più importante del Mediterraneo.
the ancient Rome was the most important city of the Mediterranean
b. Roma antica fu la città più importante del Mediterraneo.
Rome ancient was the most important city of the Mediterranean
c. *Antica Roma fu la città più importante del Mediterraneo.
ancient Rome was the most important city of the Mediterranean
(Longobardi 2005: 9-10)

(3) a. La mia casa è bella
the my home is beautiful
b. casa mia è bella
home my is beautiful
c. *mia casa è bella
my home is beautiful

Interestingly, a subset of kinship terms behaves like proper names and the item 'casa', allowing N-to-D movement, as illustrated in (4) (cf. Longobardi 1994, 1996). Still, as shown in (1), they can appear, as least for some kinship terms, determiner-less also in their base position (as in (4a)).[2]

[2] An anonymous reviewer asked how it is possible to account for the (micro)parametric variation attested in Romance. For instance, in Spanish, with possessive modification, determiners are not available independently of the presence of evaluative morphemes (e.g. (*la) mi hermana, my sister, (*el) mi gato, my cat, (*la) mi hermanita, my little sister, (*el) mi gatito, my kitty). For the sake of the present discussion, we may simply assume that Spanish possessives (like their English counterparts) are directly merged in D, blocking any other lexical item in such position.

(4) a. mio nonno è andato al mare
My grandpa went to the sea
b. nonno mio è andato al mare
my grandpa went to the sea
c. il mio nonno è andato al mare
my grandpa went to the sea

Longobardi (1996) proposes an analysis for the whole set of kinship expressions in Italian, such as *mia sorella* 'my sister' in (1a), also likely to trigger N-raising, along with additional syntactic operations. He argues for a direct link between proper names and kinship terms and he basically assumes that determiners are "expletive" when they show up with items that are prone to N-to-D movement. In subsequent papers, Longobardi connects the movement of a noun to (a phonetically empty) D with the obligatorily cooccurrence of an overt or understood genitive argument (see Longobardi 2005, 2008), assuming that such entities rigidly refer to particular individuals, from a semantic viewpoint. Recently, Giusti (2015, cf. also Cardinaletti and Giusti 2018) argues that kinship terms modified by possessive items are substantially analogous to proper names, given that they are semantically interpreted as *rigid designators* (Kripke 1980). For this reason, both nominal classes lack definite determiners in Italian.[3]

Giusti (2015), Cardinaletti and Giusti (2018) specifically propose that rigid designators project reduced syntactic nominal structures in order to address the issue of silent Ds with kinship terms (and proper names). They assume that while common nouns project three different layers (the lexical NP, the modification field including a set of functional projections FP, and the referential layer DP) rigid designators only project two layers, namely the lexical NP and the referential DP. According to Cardinaletti and Giusti (2018: 141), the possessive adjective which signals the possessor of the designator is theta-interpreted and referentially interpreted in a SpecNP position, which is immediately lower than D, given the absence of a functional layer sandwiched between D and N. In (5), we outline a sketch of their model. In (5a), the kinship term shows up with a silent D. In (5b) the proper name moves to D, following the standard analysis of Longobardi (1994). The parallelism between the two classes of nominal items is supported by the fact, already highlighted in (4b), that a subset of kinship terms can raise to D, as in (5c).[4]

[3] A close link between kinship terms and proper names is confirmed on typological grounds. For instance, Dahl and Koptjevskaja-Tamm (2001) assumes a rough hierarchy establishing that if any other kin terms are used like proper names, then those encoding ascending relations, and in particular direct ones, like 'father' and 'mother', will be used like proper names. Pham (2011) assumes that Vietnamese kinship terms have essentially the same distribution of D items (just like proper names). Stolz, Levkovych and Urdze (2017) show that in Faroese a special clitic marker (*sa*) developed that can only be suffixed to personal name and kinship possessors; if the possessor is a common noun or a place name, possession is canonically expressed via a prepositional phrase in that language (cf. also Schlücker and Ackermann 2017).

[4] Another parallel between proper names and kinship terms in Italian, as highlighted in Giusti (2015), is that they can appear without a determiner only when they are to singular. Consider the examples in (i):

(i) a. *(i) miei fratelli
the my brothers
b. *(i) Rossi
'the Rossi family'

(5) a. [$_{DP}$ Ø [$_{NP}$ mia sorella]]
"my sister"
b. [$_{DP}$ Maria [$_{NP}$ mia ~~Maria~~]]
"my Maria"
c. [$_{DP}$ mamma [$_{NP}$ mia ~~mamma~~]]
"my mother"

We assume that an analysis of this kind is essentially correct, in spite of the fact that it cannot really dispense from an *expletive* nature for the determiner in contexts where proper names are left in their base position (cf. *(*il) mio Gianni* in (1b)). Still, we have to find an explanation for the fact that kinship terms cannot behave as 'rigid designators' when they bear evaluative morphology. Namely, we have to explain why it is impossible for *evaluated* kinship terms to appear determiner-less, as illustrated above in (1c). Actually, from a semantic point of view, it is not easy to immediately catch the difference between them: evaluated and "bare" kinship terms both seem to rigidly denote individuals, semantically.

In what follows, we will try to advance a morphosyntactic analysis capable of understanding the reason why the application of evaluative affixes influences the syntactic context in which kinship terms can be employed. In section 2 we will illustrate the analysis of evaluative morphemes advanced in Savoia *et al.* 2017 (cf. also Franco, Baldi and Savoia 2020). Section 3 will apply Savoia *et al.*'s theoretical model to the issue of evaluated kinship terms. Section 4 is devoted to introduce and examine potential counterexamples and other issues relevant for our analysis. The Conclusion follows.

2. The morphosyntax of evaluatives

Following Manzini and Savoia (2017a, 2017b), for Italian/Romance, we assume a nominal morphosyntax in which the first component is a category-less root √ (cf. Marantz 1997). Next to the root √ we find different types of morphemes, including derivational and inflectional ones. Inflectional morphemes generally follow derivational suffixes. The root is interpreted as a predicate, that has one open argument place (the R-role, cf. Williams 1994), which is ultimately bound by a determiner or quantifier operator (see Higginbotham 1985). Gender and number specifications — labelled as Class — apply to the argument *x* open at the predicate. We assume that they work as predicates themselves, restricting the content of the argumental variable, eventually topped off by a determiner/quantifier.[5]

Manzini and Savoia argue that the inflectional vowel of Italian occupies an Infl(ection) position embedding the root √ and the Class nodes, which encode gender ([feminine], [masculine], etc.) or number specifications ([singular], [plural], etc.). The content of the plural is represented as the relation part-whole [⊆], specifying that the *denotatum* of the predicate can be partitioned into subsets (cf. Manzini and Savoia 2011 and subsequent literature). A rough sketch of their model is represented in (6) for the nominal items *ragazz-o/i* 'boy/boys', *ragazz-a/e* 'girl/girls'.

[5] In their framework, all lexical material is associated with interpretable contents; this proposal is not substantially different from the conception of Agree in Chomsky (2001), expressing an identity relation between features under locality (Minimal Search).

(6) a. [[*ragazz* √] [$_{\subseteq}$] [fem]/[masc] $_{\text{Class}}$] [$_{\subseteq}$] $_{\text{Infl}}$ *-a -e -i -o*]

Furthermore, following Borer (2003, 2014) and Savoia *et al.* (2017; see also Franco 2017), we take that the same lexical content can be expressed by inflectional and derivational morphemes (as well as stand-alone lexical items). For instance, in Romance, inflectional morphemes are able to introduce properties which usually are introduced by derivational tools, for example category change, size properties (e.g. *melo* 'apple tree'/*mela* 'apple'; *buco* 'hole'/*buca* 'pit', cf. Franco, Manzini and Savoia 2015; Franco, Baldi and Savoia 2020; Manzini 2020). *Vice versa*, derivational morphemes can introduce kinds of contents generally associated with inflection, as for instance gender specifications, which can be also introduced in Italian by the derivational suffix *-ess* (e.g. *maestr-o-maestr-a* 'male/female teacher' vs. *avvocat-o*, *avvocat-ess-a* 'male/ female lawyer').

For what concern evaluatives, in keeping with Savoia *et al*. (2017), we argue that they mostly express size / measure (Class) properties or the grade of physical or culturally-determined properties with reference to lexical roots, as in the examples in (7) where kinship terms bear diminutive/augmentative morphemes.

(7)	a.	sorell-in-a	/	fratell-on-e
		sister-dim-Infl		brother-aug-Infl
		'small/ younger sister'		'big/ older brother'

Evaluatives can be understood as (Class)ifying predicates that contribute to restricting the argumental variable of the root √, like gender and number layers (see also Percus 2011). Actually, it is possible to extend the analysis of adjectives in Parsons (1979) to evaluatives. For Parsons (1979: 157), adjectives in attributive positions are "operators on the predicate contributed by the noun … these operators can be further analysed in terms of conjunction with a predicate', whereby 'red box' can be translated as 'x is red and a box' ".

Given this, it is possible to follow Cinque (2010) in pursuing the idea that evaluative morphology is essentially intersective, namely it invariantly corresponds to the restrictive and intersective (or subsective) interpretation of adjectives,[6] giving rise to the interpretations in (8).

(8) a. *sorellina* λx[(sister)x & (little)x], *fratellone* λx [(brother)x & (big)x]

Syntactically, we argue that an evaluative interpretation arises from a structure in which the evaluative morphemes embed the Class node and the root, as illustrated in (9) for diminutives and augmentatives.[7] The evaluative Class item merges with the root √ combined with

[6] This interpretation is supported by the fact that diminutive morphemes can be attached to postnominal adjectives but not in pre-nominal ones in Italian, as illustrated in (i) (see Savoia *et al.* 2017).

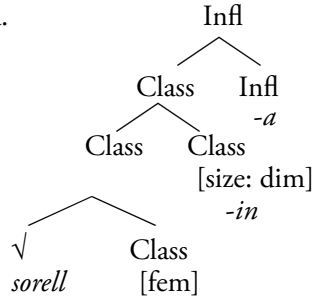
(i) a. *una grand-in-a casa/una casa grand-in-a
[a] big-dim house/a house big-dim
b. *una bell-in-a ragazza/una ragazza bell-in-a
[a] cute-dim girl /a girl cute-dim

Following Cinque (2010), pre-nominal adjectives have a non-restrictive reading (cf. also Partee 2007), while post-nominal adjectives have an intersective meaning because they restrict the denotation of the head noun with the introduction of further specifications on its range of reference.

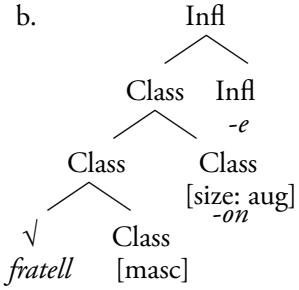
[7] In (9) Dim/Aug are introduced as [Dim]Class/[Aug]Class, and so on in the case of the other evaluative class nodes, as for instance pejorative (e.g. *sorell-acci-a*, bad sister) and endearing (e.g. *zi-ett-a*, auntie) morphemes. For the ordering and mutual exclusions between evaluative suffixes, not strictly relevant for the present discussion, see the analyses in Cinque (2015) and in Savoia *et al.* (2017).

other Class specification (e.g. gender properties) below the inflectional/agreement layer Infl.[8] We assume a nominal structure in which inflectional and derivational morphemes enable a morpho-syntactic skeleton which provides different semantic (Class) interpretations.

(9) a.



b.



3. The interaction of kinship terms with evaluative morphemes: an analysis

Given this theoretical background, in this section we will try to account for the impossible selection of determiner-less kinship terms modified by possessives, when they bear evaluative morphology. Longobardi (2005) noticed that the rigid denotation of an individual entity, from a morphosyntactic viewpoint, is a derivational property and not a lexical one. Namely, a 'rigid designation' cannot be predicted on the basis of the lexical features of a nominal item noun. It is determined through the syntactic module in a derivational fashion. Thus, proper names (and other rigid designators) are allowed to fill an empty determiner position only given certain syntactic context. For instance, Longobardi shows that in those contexts where proper names are treated like common nouns, pronouns (i.e. standard D items) are ungrammatical, as illustrated for restrictive relative modifications in (10) (cf. also Matushansky 2009):

(10) a. Il (simpatico) Gianni che conoscevo non esiste più`.
'The (nice) Gianni that I used to know no longer exists.'

[8] Other recent proposals in the generative framework aim to relate the nature and morphemic status of evaluatives to the discussion on the internal structure of the noun, also in a cross-linguistic perspective, see, among others, Wiltschko (2006), Ott (2011), de Belder, Faust and Lampitelli (2014), Cinque (2007, 2015), Franco, Baldi and Savoia (2020). For instance, Wiltschko (2006) assumes that diminutive suffixes are 'light nouns', and specifically that they correspond to numeral classifiers. Ott (2011) proposes a detailed structure where the diminutives are analysed as 'numeral classifiers'. Both authors converge in identifying the diminutive suffix with a lexicalization of a nominal layer independent of the root.

b. *Gianni (simpatico) che conoscevo non esiste più`.
'Gianni (nice) that I used to know no longer exists.'
c. *Il (simpatico) lui che conoscevo non esiste più`.
'The (nice) he that I used to know no longer exists.'
d. *Lui (simpatico) che conoscevo non esiste più`.
'He (nice) that I used to know no longer exists.'

(Longobardi 2005: 16)

Crucially, when kinship terms are modified by intersective/subsective adjectives, they cannot appear determiner-less, as in (11) where the adjective *imbranata* 'clumsy' modifies the kinship term for sister.

(11) a. *(la) mia sorella imbranata ha cambiato strada
the my sister clumsy has changed road
'my clumsy sister changed direction'

In the terms of Cardinaletti and Giusti (2018), if modifiers enter the derivation, the kinship term is no more a rigid designator and works as the head noun of a (standard) syntactic structure in which the functional field is projected. In such (canonical) structure the possessor moves from SpecNP to SpecPossP, roughly as illustrated in (12).

(12) a. [$_{DP}$ la [$_{PossP}$ mia [$_{FP}$ sorella [$_{FP}$ [$_{AP}$ imbranata] … [$_{NP}$ ~~mia sorella~~]]]]]
the my sister clumsy
'my clumsy sister'

Thus, both Longobardi and Cardinaletti and Giusti assume that it is impossible to select a lexical item as a rigid designator when it undergoes subsective/intersective/restrictive modification. Nevertheless, Cardinaletti and Giusti assumes that rigid designation is a lexical property of nominal items, while Longobardi claims that rigid designation arises derivationally, given certain syntactic environments.

As we have seen in the previous section, Savoia *et al.* (2017) argue that it is possible to treat evalutative morphemes as classifying morphemes that operate a [Size] restriction of the lexical root. This is a morphological operation, which is however visible at the level of syntax. Indeed, diminutives and augmentatives, as in (7)-(8), seem to work just like intersective adjectives, as in (13). This clearly blocks the possibility for kinship terms to act as rigid designators which allow an empty D position possibly filled via N to D movement.

(13) a. la mia sorellina ≈ la mia sorella piccola
the my sister.dim the my sister little
'my little sister/my younger sister'
b. il mio fratellone ≈ il mio fratello grande
the my brother.aug the my brother big
'my big brother/my older brother'

This pattern extends to all the other kinds of evaluatives in Italian. Pejoratives are a case in point, given that they seem to show that Longobardi is right in assuming that 'rigid designation' is shaped 'compositionally' by the morphosyntactic derivation (cf. also Baggio and Cairncross 2020 for a recent cartographic proposal which supports this view) and cannot be predicted

on the basis of the lexical features of a given nominal item. Consider for instance the case of the Italian pejorative suffix *-astr-* when it applies to kinship terms, as in the examples in (14). These lexical items, from a semantic viewpoint, seem to lexically work as rigid denotators, given that they encode particular specimen of the class of kinship terms. In (14), they can denote a daughter/son of one's step-parent by a marriage other than with one's own father or mother. In these contexts, it is not possible to retrieve a [size] interpretation for these evaluative morphemes and they are *prima facie* connected to a *denotatum* which is fully stored in the lexicon (i.e. a fossilized form).[9]

(14)	a.	sorell-astr-a	b.	fratell-astr-o	
		step-sister, half-sister		step-brother, half-brother	

Actually, such items are fully ungrammatical without a definite determined in the possessor construction we have outlined so far, as illustrated in (15), just like the examples we have provided for kinship terms bearing diminutive or augmentative affixes.

(15)	a.	*(la)	mia	sorellastra /	*(il)	mio	fratellastro
		The	my	step/half sister	/ The	my	step/half brother

Thus, it seems that the morpheme *-astr* is still active from a morphosyntactic viewpoint, given that it behaves just like the other evaluatives introduced in the discussion so far. The Italian suffix *-astr* derives from the Latin morpheme *-aster*, which encodes resemblance/similarity (cf. Thomas 1940; Rohlfs 1968). The evaluative (pejorative) meaning of this suffix, according to Merlini Barbaresi (2004), comes from the fact that the nominal items bearing this morpheme are all characterized by qualities which are *similar* to those of the respective bases, but less precise, less perfect.[10] Consider for instance the examples in (16a) where the lexical items have a clear pejorative flavour and the items in (16b) in which they encode an approximate (i.e. less intense, less clear, etc.) version of the lexical base.

(16)	a.	poet-astr-o	>	poeta,	filosof-astr-o	>	filosofo
		bad poet		poet	bad philosopher		philosopher
	b.	ross-astr-o	>	rosso	sord-astr-o	>	sordo
		reddish		red	slightly deaf		deaf

We argue that the [approximate] value of the items in (16b) is shared by the kinship terms bearing the *-astr* suffix.[11] Thus, it is arguable to propose, for these nominal elements, an

[9] Indeed, Merlini Barbaresi (2004: 292) precisely assumes that items like *fratellastro* 'step-brother, half-brother', *figliastro* 'step-son', *sorellastra* 'step-sister, half-sister' are fully lexicalized in Italian.

[10] Serianni (1988: 550) points out that the suffix *-astr* can also convey an endearing or playful meaning, as with items like *topastro* 'little funny rat', *giovinastro* 'youngster'.

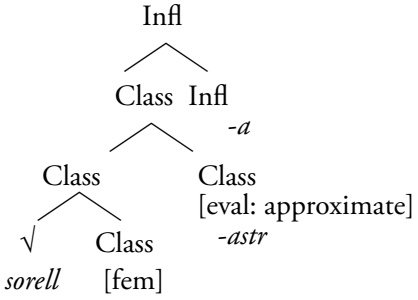
[11] We can propose a similar 'approximate' value for the the (scarcely productive) evaluative suffix *-ign*, also employed with kinship terms, as in (i), which disallow an empty D as in (ii). The approximate (i.e less intense) value of this suffix is found with items like the one in (iii).

(i)	patr-ign-o,		matr-ign-a				
	'stepfather'		'stepmother'				
(ii)	*(la)	mia	matrigna /	*(il)	mio	patrigno
	The	my	stepmother /	The	my	stepfather

intersective/subsective semantics like the one represented in (17) and a morphosyntax like the one sketched in (18).

(17) a. *Sorellastra* λx[(sister)x & (approximate)x], *fratellastro* λx [(brother)x & (approximate)x]

(18) a.



We argue that not only the outer node (Infl) of the nominal structure is recruited for the syntactic computation (e.g. for Agree operations), but also the Class layers of the nominal skeleton acting as predicates which restrict the referential properties of the root. In this precise context, they turn kinship terms, which are a set of nominal items sharing many features with proper names, into common nouns by applying a functional restriction on their reference. This, in turn, blocks the availability of an empty D position for kinship items. The fact that derivational morphemes of this kind alter the syntactic environment in which a noun is couched support the theory of grammar advanced by Manzini and Savoia (2007, 2011), who precisely assume that Merge takes morphemes as its input and single morphemes are visible to syntactic computation (in the terms of Cardinaletti and Giusti 2018, Class morphemes feed the functional field sandwiched between N and D, blocking rigid designation).

4. Further issues on possessives, kinship terms and evaluatives

In the previous section, we have sketched the main point of our analysis which addressed the interaction between evaluative morphology and kinship terms introduced by possessives in Italian, showing how the application of evaluative affixes influences the syntactic context in which kinship terms can be employed, blocking the availability of an unpronounced definite D. Now, we address some facts that can be considered as potential problems for our claims.

First, we have said that we are assuming that morphological (e,g. evaluatives) and syntactic (e.g. adjectives) modifiers block rigid denotation for kinship terms in the spirit of Cardinaletti and Giusti (2018). Actually, there is a potential counterexample involving the comparative (suppletive) form of the adjectives which exactly identifies [size] properties in Italian, namely ʻminoreʼ (smaller, younger) and ʻmaggioreʼ (bigger, older). Consider the examples in (19).

(19) a. Mia sorella minore è andata al parco
'My younger sister went to the park'

(iii) aspr-ign-o *from* aspro
slightly sour sour

b. Mio fratello maggiore è andato al parco
'My older brother went to the park'

In (19) we find that the kinship terms in the possessor construction have a modifier and still allow an empty D, *contra* what would be expected. We argue that items like 'sorella minore' and 'fratello maggiore' in (19) are treated like lexical compounds and cannot be syntactically rearranged. That is the reason why they are treated as lexical members of the class of kinship terms and can still work as rigid designators. This is confirmed by the fact that these constructions cannot be modified by intervening intensifiers or by other lexical material, as in (20).

(20)	a. *un	fratello	molto	minore	(cf. e.g. un percorso molto minore)
	a	brother	much	younger	'a much shorter path'
	b. *una	sorella	ben	maggiore	(cf. e.g. una fatica ben maggiore)
	a	sister	much	older	'a far greater effort'

Furthermore, the items 'minore' and 'maggiore' are commonly recruited in Italian to form various other lexical compounds/multiword expressions, as shown in (21).

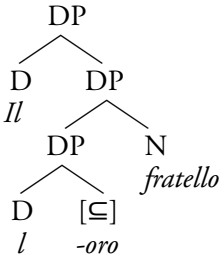
(21)	Orsa maggiore	panda minore	Si maggiore	…
	Ursa mayor	red panda	B major	

A further note concerns the paradigm of possessives in front of kinship terms in Italian. Consider the data in (22)

(22)	sing			plur	
	1st	mio padre	'my father'	nostro padre	'our father'
	2nd	tuo padre	'your father'	vostro padre	'your father'
	3rd	suo padre	'his/her father'	??(il) loro padre	'their father'

According to our judgement, the third person plural possessive *loro* disallows the possibility of a bare determiner in the possessor – kinship term structure. The special properties of the pronoun *loro* have attracted great interest in the literature, starting from the work of Cardinaletti (1998), Cardinaletti and Starke (1999) concerning a fine grained typological of *strong* vs. *weak* pronouns in Italian. Here we do not enter into technical details and we just follow Manzini (2014; cf. also Manzini and Savoia 2015; Baldi and Savoia 2019, 2021) in assuming that 'loro' includes an inflectional oblique morpheme *-oro*, notated as the (possession) relation [⊆], taking, in genitive contexts, the pronominal/definiteness base *l-* to which it attaches as its internal argument and the head noun as its external argument. A possible representation is sketched in (23).

(23) a.



The other possessive items, introducing the participants in the speech act (1st/2nd person, ‘mio/nostro’, ‘tuo/vostro’) or the singular reference to a discourse anaphoric argument (3rd singular person *suo*) do not require the presence of a definite determiner, possibly because their interpretive content is able to exhaustively externalize the referential domain of DP.

Our sketchy idea is that the definiteness *l-* base (the same base found in Italian definite determiners, e.g. *il*, the.m.sg, *la*, the.f.sg, etc.) encoded in the pronoun ‘loro’ triggers the spell-out of a higher D item topping off the structure, given the argumental nature of the lower D item in ‘loro’. Such lower argumental D requires an (overt) higher D fixing its referential properties (cf. Lekakou and Szendrői 2012; Franco, Manzini and Savoia 2015b). It is possible, in other words, that the lower D simply values the argument slot of [⊆], awaiting further quantificational closure (operated by the higher D). We leave this matter for future works.

5. Conclusion

In this brief paper we have addressed the interaction between evaluative morphology with kinship terms introduced by possessives in Italian, showing that the application of evaluative affixes influences the syntactic context in which kinship terms can be employed. We have argued that this (trivial) empirical observation has some interesting consequences from a theoretical viewpoint: the fact that derivational morphemes, such as evaluatives, alter the syntactic environment in which a noun is couched support the theory of grammar advanced by Manzini and Savoia (2007, 2011), who assume that Merge takes morphemes as its input and single morphemes are visible to syntactic computation.

Specifically, we have argued that it is possible to treat evalutative morphemes as classifying morphemes that operate a [Size/Approximation, etc.] restriction of the lexical root. This is a morphological restriction, which is however visible at the level of syntax. Indeed, diminutives and augmentatives seem to work just like intersective/subsective adjectives activating a functional skeleton sandwiched between the determiner and the noun. This blocks the possibility for kinship terms in possessive contexts to act as rigid designators allowing an empty D position, potentially filled via N to D movement.

References

Baggio, Pietro, and Alexander Cairncross. 2020. "The Syntax of Puss in Boots and Sleeping Beauty. Proper Names between Productivity and Atomicity." In *Proceedings of ConSOLE XXVIII*, ed. by Astrid van Alem, Mirella De Sisto, Elisabeth J. Kerr, and Joanna Wall, 51-74. Leiden: Leiden University Centre for Linguistics.

Baldi, Benedetta, and Leonardo M. Savoia. 2019. "Possessives in North-Calabrian Dialects." *Quaderni del Dipartimento di Linguistica* 27: 19-40.

Baldi, Benedetta, and Leonardo M. Savoia. 2021. "Possessives, from Franco-Provencal and Occitan Systems to Contact Dialects in Apulia and Calabria." *Languages* 6 (63): 1-28.

Borer, Hagit. 2003. "Exo-skeletal vs Endo-skeletal Explanations: Syntactic Projections and the Lexicon." In *The Nature of Explanation in Linguistics Theory,* ed. by John Moore, and Maria Polinsky, 31-67. Stanford, CA: CLSI Publications.

Borer, Hagit. 2014. "Derived Nominals and the Domain of Content." *Lingua* 141: 71-96.

Cardinaletti, Anna. 1998. "On the Deficient/Strong Opposition in Possessive Systems." In *Possessors, Predicates and Movement in the Determiner Phrase*, ed. by Artemis Alexiadou, and Chris Wilder, 17-53. Amsterdam: John Benjamins.

Cardinaletti, Anna, and Michal Starke. 1999. "The Typology of Structural Deficiency. A Case Study of the Three Classes of Pronouns." In *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 145-233. Berlin: Mouton-De Gruyter.

Cardinaletti, Anna, and Giuliana Giusti. 2018. "Micro-Variation in the Possessive Systems of Italian Dialects. In Language Use and Linguistic Structure." In *Proceedings of the Olomouc Linguistics Colloquium*, ed. by Joseph Emonds, Markéta Janebová, and Ludmila Veselovská, 137-154. Olomouc: Palacký UP.

Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: A Life in Language,* ed. by Michael Kenstowicz, 1-54. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2007. "La natura grammaticale del diminutivo e del vezzeggiativo." In *Miscellanea di studi linguistici offerti a Laura Vanelli da amici e allievi padovani*, ed. by Roberta Maschi, Nicoletta Penello, and Piera Rizzolatti, 229-236. Udine: Università degli Studi di Udine.

Cinque, Guglielmo. 2010. *The Syntax of Adjectives: A Comparative Study.* Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2015. "Augmentative, Pejorative, Diminutive and Endearing Heads in the Extended Nominal Projection." In *Structures, Strategies and Beyond*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 67-81. Amsterdam: John Benjamins.

Dahl Östen, and Maria Koptjevskaja-Tamm. 2001. "Kinship in Grammar." In *Dimensions of Possession*, ed by. Irène Baron, Michael Herslund, and Finn Sørensen, 201-243. Amsterdam: John Benjamins.

De Belder, Marijke, Noam Faust, and Lampitelli, Nicola. 2014. "On a Low and a High Diminutive." In *The Syntax of Roots and the Roots of Syntax,* ed. by Alexiadou, Artemis, Hagit Borer, and Florian Schäfer, 149-163. Oxford: Oxford UP.

Franco, Ludovico, M. Rita Manzini, and Leonardo M. Savoia. 2015a. "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.

Franco, Ludovico, M. Rita Manzini, and Leonardo M. Savoia. 2015b. "Linkers and Agreement." *The Linguistic Review* 32: 277-332.

Franco, Ludovico, Benedetta Baldi, and Leonardo M. Savoia. 2020. "Collectivizers in Italian (and Beyond). The Interplay between Collectivizing and Evaluating Morphology (and the Div paradox)." *Studia Linguistica* 74: 2-41.

Franco, Ludovico. 2017. "Some Notes on Italian Ethnic Adjectives." *QULSO* 3: 87-116.

Giusti, Giuliana. 2015. *Nominal Syntax at the Interfaces. A Comparative Study of Languages with Articles*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Higginbotham, James. 1985. "On Semantics." *Linguistic Inquiry* 16: 547-621.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard UP.

Lekakou, Marika, and Krista Szendrői. 2012. "Polydefinites in Greek: Ellipsis, Close Apposition and Expletive Determiners." *Journal of Linguistics* 48: 107-149.

Longobardi, Giuseppe. 1994. "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form." *Linguistic Inquiry* 25: 609-665.
Longobardi, Giuseppe. 1996. *The Syntax of N-Raising: A Minimalist Theory*. Utrecht: OTS Working Papers.
Longobardi, Giuseppe. 2005. "Toward a Unified Grammar of Reference." *Zeitschrift für Sprachwissenschaft* 24: 5-44.
Longobardi, Giuseppe. 2008. "Reference to Individuals, Person, and the Variety of Mapping Parameters." In *Essays on Nominal Determination: from Morphology to Discourse*, ed. by Alex Klinge, and Henrik Høeg Müller, 189-211. Amsterdam: John Benjamins.
Manzini, M. Rita. 2014. "Grammatical Categories: Strong and Weak Pronouns in Romance." *Lingua* 150: 171-201.
Manzini, M. Rita, and Leonardo M. Savoia. 2007. *A Unification of Morphology and Syntax. Investigations into Romance and Albanian Dialects.* London: Routledge.
Manzini, M. Rita, and Leonardo M. Savoia. 2011. *Grammatical Categories: Variation in Romance Languages*. Cambridge: Cambridge UP.
Manzini, Maria Rita, and Leonardo M. Savoia. 2017a. "Gender, Number and Inflectional Class in Romance: Feminine/plural *–a*." In *Language Use and Linguistic Structure*, ed. by Joseph Emonds, and Markéta Janebova, 263-282. Olomouc: Palacky University Olomouc.
Manzini, M. Rita, and Leonardo M. Savoia. 2017b. "N Morphology and Its Interpretation: The Neuter in Italian and Albanian Varieties." In *Constraints on Structure and Derivation in Syntax*, *Phonology and Morphology*, ed. by Anna Bloch-Rozmej, and Anna Bondaruk, 213-236. Frankfurt am Main: Peter Lang.
Manzini, M. Rita. 2020. "The Morphosyntactic Structure of Number in Italian and Albanian. High and low plurals." *Catalan Journal of Linguistics* 19: 127-157.
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.
Masi, Sonia. 2020. "Variation in the Possession of Kinship Terms in the Dialects of Italy." *Proceedings of the 2020 CLA - Canadian Linguistic Association meeting.*Matushansky, Ora. 2009. "On the Linguistic Complexity of Proper Names." *Linguistics and Philosophy* 31: 573-627.
Merlini Barbaresi, Lavinia. 2004. "Alterazione." In *La formazione delle parole in italiano*, ed. by Maria Grossmann, and Franz Rainer, 264-292. Tübingen: Niemeyer.
Ott, Dennis. 2011. "Diminutive-formation in German. Spelling out the Classifier Analysis." *The Journal of Comparative Germanic Linguistics* 14: 1-46.
Parsons, Terence. 1979. "An Analysis of Mass Terms and Amount Terms." In *Mass Terms: Some Philosophical Problems*, ed. by Francis J. Pelletier, 137-166, Dordrecht: Springer.
Partee, Barbara H. 2007. "Compositionality and Coercion in Semantics: The Dynamics of Adjective Meaning." In *Cognitive Foundations of Interpretation*, ed. by Gerlof Bouma, Irene Kramer, and Joost Zwarts, 145-161. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
Percus, Orin. 2011. "Gender Features and Interpretation: a Case Study." *Morphology* 21: 167-196.
Pham, Mike. 2011. "Are Vietnamese Kinship Terms Pronouns?" Ms. University of ChicagoRohlfs, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia*. Torino: Einaudi.
Savoia, Leonardo M., M. Rita Manzini, Benedetta Baldi, and Ludovico Franco. 2017. "A Morphosyntactic Analysis of Evaluatives in Italian." *SILTA - Studi Italiani di Linguistica Teorica e Applicata* XLVI: 413-440.
Schlücker, Barbara, and Tanja Ackermann. 2017. "The Morphosyntax of Paper Names: an Overview." *Folia Linguistica* 51: 309-333.
Serianni, Luca. 1988. *Grammatica italiana. Italiano comune e lingua letteraria. Suoni, forme, costrutti*. Torino: Utet.
Stolz, Thomas, Nataliya Levkovych, and Aina Urdze. 2017. "Die Grammatik der Toponyme als typologisches Forschungsfeld: eine Pilotstudie". In *Namengrammatik*, ed. by Johannes Helmbrecht, Damaris Nübling, and Barbara Schlücker, 121-146. Hamburg: Buske.
Thomas, François. 1940. "Le suffixe Latin -aster/ -astrum." *Revue des Études Anciennes* 48: 520-528.

Williams, Edwin. 1994. *Thematic Structure in Syntax.* Cambridge, MA: MIT Press.
Wiltschko, Martina. 2006. "Why Should Diminutives Count?" In *Organizing Grammar. Linguistic Studies in Honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Riny Huijbregts, Ursula Kleinhenz, and Jan Koster, 669-679. Berlin: Walter de Gruyter.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: M.R. Manzini (2021) Chomsky's (2020) Links and linker phenomena. *Qulso* 7: pp. 89-102. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12004

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Chomsky's (2020) Links and linker phenomena

M. Rita Manzini
Università degli Studi di Firenze (<mariarita.manzini@unifi.it>)

Abstract:

In Section 1, I present a recent formalisation of modification and conjunction structures by Chomsky (2020), in terms of sequences of Pair-Merge units, each including a Link element identified with the nominaliser/verbaliser *n/v*. In Sections 2-3, I argue that the Link element is overtly visible in many languages in the form of (adnominal) linkers. In Section 4, I suggest that Case stacking, and Concord may also be viewed as externalizations of Links.

Keywords: *Adjectives, Linkers, Pair-Merge, Possessives*

1. Introduction: coordination and modification according to Chomsky (2020)

Chomsky (2020: 49-52), addressing adjectival (adnominal) modification, considers examples like (1a). He remarks that "there's an asymmetry between the two words, that's clear. The element that's formed is a noun phrase, not an adjective phrase. So *young* is an adjunct that's not changing the category". As for (1b), the issue is even worse, given that the Labeling Algorithm of Chomsky (2013) predicts that in X-YP mergers, it is the head X that provides the label. "*Portrait of John* is not a head, and could be arbitrarily complex, but it still provides the label of the whole unit. So there is a clear asymmetry".

(1) a. young man
b. old portrait of John

For Chomsky, the way to a different analysis passes through "unbounded, unstructured coordinations" like (2). His idea is that "at some point in [the] generation of an interpretable structure, each of the adjective phrases in [2] will be predicated of *someone* with an asymmetric relation similar to *young man*,

presumably pair-formation". To formalize this relation "it seems we need an operation Pair-Merge, which will also apply to the simple adjunct case like *young man*". Importantly, for the purposes of labeling – or more generally of projecting the structure, "*young* will be adjoined to *man*, but you don't see it in the labeling because it's off in some other dimension […] It's not two dimensional like a blackboard, but there's no reason to suppose that mental representations are restricted this way".[1]

(2) I met someone young, happy, eager to go to college, tired of wasting his time, …

The next step is acknowledging that the structure formed by the modifiers in (1) is "not just a set of paired elements, it's a sequence of paired elements". The argument is that the order of the unstructured unbounded elements matters because of reasons that were pointed out by Jim McCawley over half a century ago, namely, notions like *respectively*. So in (3a) the order in which the coordinated elements appear "affects the semantic interpretation". The same is true of adjuncts, as in (3b).

(3) a. John and Bill saw Tom and Mary, respectively
b. John and Bill are young and tall, respectively

Formally, "you generate syntactic objects in WS [Work Space], select a finite set of these, and from that set form a sequence S… and that sequence S is the syntactic object that you're then going to merge into the construction". As illustrated in (4), "we're forming a sequence which begins with some conjunction, and then contains elements, each of which is predicated of something. So we have a sequence of elements that looks like [4], with links L_i". Of course, "when you get down to the limiting case, when n = 1, that's just plain adjunction. So *the young man*".

(4) $< \text{CONJ}, < S_1, L_1>, \ldots ,< S_n, L_n>>$

A number of interesting consequences follow from this treatment of modification and conjunction. Specifically,

> if you have the phrase *old man*, you can't extract *man* and leave *old*; you can't extract *old* and leave *man*. So the elements of the pairs are inaccessible […]. You have the coordinate structure constraint because every term is inaccessible. You have the adjunct island constraint because you can't pull the elements out of adjuncts.

What I am interested in is the question, "what is L? What do things link to?" The suggestion is that "L for nominal sequences is just *n*, the categorizer of each of the coordinated phrases". As Chomsky points out, the basic nature of *n* and *v* is that of nominal and verbal categorizers, assuming that roots are themselves deprived of a category (Marantz 1997). Therefore, "the sequence S is actually a Root Phrase (RP) sequence which is categorized by linking to *n*, *v*, *a*, in the basic cases".

[1] In fact, according to Chomsky (*ibidem*) "the unbounded unstructured cases like [2] show you in effect that there are unboundedly many dimensions as adjective phrases are independently adjoined to the host".

In a later passage, Chomsky (2020: 56) discusses a model of head raising by Pair-Merge attributed to H. Kitahara (see Epstein, Kitahara and Seely 2016). T-to-C is the instance of head raising being considered and the result of the derivation is the structure in (5). The derivational steps listed in (6) provide a useful sketch of how Set-Merge and Pair-Merge are interwoven in the derivation.

(5) {<C, T> {T, VP}}

(6) a. you generate {T, VP}.
b. Then you pair-merge C and T, yielding <C, T>.
c. The workspace now contains <C, T> and {T, VP}.
d. you now merge <C, T> to {T, VP}, yielding {<C, T> {T, VP}}.

My topic here is modification, especially adnominal modification (hence neither coordination, nor head raising *per se*). I will address a very narrow question, though one with potentially wide-ranging empirical consequences. In its general form, the question is whether one might see a morphological realization of the Link in modification structures like (4). More specifically, a possible candidate is what is known in the literature as a linker (whether the assonance is intended by Chomsky, I don't know). Linkers include the pre-adjectival and pre-genitival articles of Balkan languages (Greek, Albanian, Romanian, Aromanian), on which I base my discussion in sections 2-3, given the relatively familiar nature of the constructs and of the morphology involved. In section 4, I briefly turn to a number of descriptive and theoretical issues that are raised by the present discussion of linkers, which I leave largely open for future research.

2. Linkers

In many languages, belonging to diverse families, direct modification of an N by adjectives, PPs (oblique case NPs) or relative clauses is not possible. Rather it is necessary to introduce the modifier by means of a functional element which we may agree to generically call a linker. These languages are therefore different from English (1a) or (2) where *man* or *somebody* are directly merged with *young* (pre-nominally or post-nominally), and the same in fact holds of PP (oblique Case) modifiers, for instance the genitive PP *of John* in (1b).

In Albanian, modification by an adjective or genitive PP implies the presence of an extra element with which the adjective or PP is merged, namely the linker. The syntactic object so created is then merged with the noun it modifies. Some examples of adjectival modification in the nominative case are provided in (7) for the masculine and in (8) for the feminine. The inflections on 'boy' and 'girl' are sensitive not only to Φ features and Case, but also to whether N is indefinite, as in the (a, b) examples, or definite, as in the (c, d) examples. This phenomenon (also found for example in Romanian) is traditionally analyzed as involving a post-nominal definite article, often derived by N to D head movement by generative scholars (Turano 2002 on Albanian). I follow Manzini and Savoia (2011, 2018) in assuming that definiteness is a feature of Albanian Ns, realized by Ninflections, like Φ features and Case (see also Chomsky 2020: 51).

(7) a. një dialë i bukur
a boy LKR.M.SG good-looking
'a good-looking boy'
b. disa djem të bukur
some boy.PL LKR.PL good-looking
'some good-looking boys'

c. djali i bukur
boy-M.SG.DEF LKR.M.SG good-looking
'the good-looking boy'

d. djemtë të bukur
boys-PL.DEF LKR.PL good-looking
'the good-looking boys'

(8) a. një vajzë e bukura
a girl LKR.F.SG good-looking-F
'a good-looking girl'

b. disa vajza të bukura
some girl.PL LKR.PL good-looking-F
'some good-looking girls'

c. vajza e bukura
girl-F.SG.DEF LKR.F.SG good-looking-F
'the good-looking girl'

d. vajzat të bukura
girl-PL.DEF LKR.PL good-looking-F
'the good-looking girls'

(7)-(8) show that pre-adjectival linkers are sensitive to the Φ features of N. The contrast between singular and plural is seen in (7a) vs. (7b) and (7c) vs. (7d) and similarly in (8). As for gender, compare (7a) to (8a), (7c) to (8c). Further data, illustrating sensitivity to definiteness and Case, are available in the formal literature (Campos 2009, Manzini and Savoia 2011, 2018, Franco, Manzini and Savoia 2015).

Linkers have nothing to do with adjectives *per se*, but rather are devices that enable adnominal modification. Thus, they are obligatorily merged with a genitive NP before the resulting syntactic object can modify N. The genitive NP can be definite or indefinite, as can be the head N. In (9) I exemplify the paradigm for nominative singular heads; all data here/and above are from Turano (2004).

(9) a. një mur i shtëpisë/ një shtëpje
a wall.M.SG LKR.M.SG house-F.SG.OBL.DEF/ a house-F.SG.OBL
'a wall of the/a house'

b. një dhomë e shtëpisë/ një shtëpje
a room.F.SG LKR.F.SG house-F.SG.OBL.DEF/ a house-F.SG.OBL
'a room of the/a house'

c. muri i shtëpisë/ një shtëpje
wall-M.SG.DEF LKR.M.SG house-F.SG.OBL.DEF/ a house-F.SG.OBL
'the wall of the/a house'

d. dhoma e shtëpisë/ një shtëpje
room- M.F.DEF LKR.F.SG house-F.SG.OBL.DEF/ a house-F.SG.OBL
'the room of the/a house'

The examples in (9) highlight a further important property of linkers in Albanian. The linker is sensitive to the Φ features, definiteness and Case of the modified N but not of the modifier NP. Thus in (9a) vs. (9b) or (9c) vs. (9d) the linker varies according to the gender of

the head noun. In (9a) and (9c), where the genders of the head N and of the genitive NP are mismatched, the linker reflects the gender of the head N (masculine). Similar observations hold for all other relevant features, as again discussed by the literature (especially Franco, Manzini and Savoia 2015, Manzini and Savoia 2018).

So far, then, I have highlighted the fact that the Albanian linker, whatever its nature may be, co-varies with the inflection of the modified N. In fact, a stronger generalization emerges, namely that the linker has (roughly) the same morphological shape as the definite nominal inflection. See for instance the nominative masculine singular (-)*i* definite inflection and linker in (7c) or the plural (-)*t(ë)* definite inflection and linker in (7d), (8d). This generalization motivates the traditional label of linkers as articles; that they are not articles, or rather determiners, in the sense of referential operators is quite straightforwardly indicated by the fact that they combine with the indefinite article in various examples in (7)-(9) (see Manzini and Savoia 2011, 2018 for further examples).

Familiar languages like Romanian, Greek also have article-like linkers. In Greek the relevant phenomenon is restricted to definite head nouns and is mostly studied under the label of polydefiniteness (Campos and Stavrou 2004, Lekakou and Szendrői 2012, Guardiano and Stavrou 2014). In Eastern Romance the phenomenon is also limited to definite Ns and more robust in Aromanian (Campos 2005, Manzini and Savoia 2018) than in Romanian. In the Iranian languages, linkers have similar distribution to Balkan linkers, in front of adjectives and oblique modifiers, though the best known of them, namely the Persian ezafe is an invariable form (Larson and Yamakido 2008). Other Western Iranian languages display the familiar patterns of sensitivity of linkers to Φ features, definiteness and Case of the modified N (e.g. Hawrami Kurdish, Holmberg and Odden 2008; Kurmanji Kurdish, Franco, Manzini and Savoia 2015). The continuity of these various phenomena, concealed by traditional labels is endorsed by typological studies, see especially Plank (1995), also for the discussion of different language families.

In short, Albanian provides a robust example of a phenomenon which is widespread in the Indo-European family, whereby nominal modification by direct merge of N(P) with an adjective or an oblique NP (Possessor) is impossible, unlike in English. Rather merger requires the presence of a linker, which is either an invariable element or an element with the same morphology as nominal (D) inflections. Given the co-occurrence of modification and coordination in Chomsky's (2020) discussion of Links, we may wonder how linkers behave under coordination. The linker can be repeated in every member of the coordination, and in Albanian it must, as in (10) (from Campos 2009: 1015, 1017).

(10) a. Vajza e mire (dhe) e sjellshme studion shumë
girl-DEF LKR good and LKR well behaved studies much
'The good (and) well behaved girl studies a lot'

b. Studentja më e urtë e klasës është Linda
student-DEF most LKR quiet LKR class-OBL.DEF is Linda
'The quietest student of the class is Linda'

Albanian linkers are also obligatory in post-copular position, as in (11), though in this respect, as for every other property discussed so far, a certain range of crosslinguistic variation is observed.

(11) a. Vajza është e bukura
girl-F.SG.DEF is LKR.F.SG good-looking-F
'The girl is good-looking'

b.	Ky	është	i	djalit
	this.M.SG	is	LKR.M.SG	boy-M.SG.OBL.DEF
	'This is of the boy'			

Before I consider the question that I set myself, namely whether linkers are exponents of the syntactic objects that Chomsky (2020) calls Links I will first review the main analyses of linkers available within formal approaches; further details can be found in Franco, Manzini and Savoia (2015). One approach, notably endorsed by Richards (2010), construes linkers (e.g. the Persian ezafe) as means for identity avoidance under adjacency. Thus occurrences of N-N and of N-A are avoided by inserting the linker between the two +N elements. This approach is called into question by the obligatoriness of linkers in post-copular (predicative) position, as in (11), where N/A is adjacent to V. Another approach rejected here is Larson and Yamakido's (2008), treating the linker as a Case marker (also for adjectives). It is perhaps not unnatural to propose this role for the ezafe in Persian, a language lacking other Case morphology. Yet it is difficult to see how this could be extended to robustly Case inflected languages like Albanian, where the linker clearly reflects the case of the modified noun, not of the modifier.

Useful insights are contained in the approach of Den Dikken (2006), Den Dikken and Singhapreecha (2004), who propose that linkers are copulas (see Campos and Stavrou 2005 for an application to Balkan languages). This brings to the fore an important conceptual theme, namely the fact that linkers involve predication environments. Yet the predicate-like treatment of the linker is once again undermined by the fact that linkers are found in post-copular contexts like (11) where a predicate is independently present (though see the discussion of (19)). More generally, linkers do not at all have verbal-like morphology but are rather nominal-like. Philip (2012) indeed raises the issue of the connection between linkers and agreement. She concludes that they are essentially Agr heads. Some technical difficulties with this proposal arise because of Chomsky's (1995, 2000) arguments against Agr(P). Leaving this aside, it seems to me that the real question is why an extra Agr element would be present, especially considering that adjectives already have an agreement inflection. For instance, *bukur* 'good-looking-M' in (7) contrasts with *bukura* 'good-looking-F' in (8).

Finally, some authors take at face value the fact that linkers, at least in Indoeuropean (I-E) languages, share the same morphology as determiners and/or pronominal clitics. Lekakou and Szendrői (2012) ultimately endorse a slightly different categorisation for D (the linker) and Def (the determiner), somewhat defeating the purpose of explaining one on the basis of the other. Franco, Manzini and Savoia (2015), Manzini and Savoia (2018) propose that linkers are Ds. As for the reason why such elements would be generated, they resort to an interface explanation, namely that in linker languages the subject of the predication needs to be represented within the maximal projection of the predicate. For instance, in (11a) the linker provides a partial saturation of the argument of 'good' within the AP projection, prior to saturation by 'the girl'. The same is true in (11b), assuming that the genitive is a two-place relation (Manzini and Franco 2016) where the internal argument is satisfied by the genitive NP itself and the external argument is the modified NP. In (11b), then, the linker provides a local saturation of the external argument of the genitive relator prior to saturation by 'this'. The problem is that it is not obvious why the relevant constraint (here local satisfaction of the valence) would hold.

In the next section, I will consider whether linkers, as illustrated in this section, can and should be analysed as exponents of Chomsky's (2020) Links, as introduced in section 1. This means answering two questions. First, do Links provide an adequate model for the properties of linkers reviewed above? Second, does an analysis based on Links, hence on Pair-Merge sequences, improve our understanding with respect to the various analyses reviewed?

3. Linkers and Links

As already mentioned, Albanian has been chosen as a language of exemplification because of the richness and regularity of the linkers phenomenon. Not all of the properties of Albanian that I have listed occur in all linkers languages, but subsets of them typically do. The first important property is that linkers in adnominal modification are nominal; concretely, they are neither prepositions/Case markers nor copulas, but rather article- or clitic-like. From a theoretical point of view, if we want to identify linkers with Links in Pair-Merge sequences, the question is whether it is plausible to categorise them as *n*.

In order to answer this question, I need to take a detour into the nature of *n*. Chomsky simply identifies *n* with the nominal categoriser and phase head. In Distributed Morphology (DM), *n* is identified with the lowest of the features of the nominal root R, namely gender in IE languages, or more generally nominal class (see Kramer 2015 for an overview). I-E nouns have a tripartite morphological structure, consisting of a root followed by a thematic vowel which generally conveys nominal and inflectional class (and sometimes number) and a higher slot associated with number, Case, definiteness (Halle and Vaux 1998). In Albanian, the three constituents can be clearly seen for instance in the (regular) definite plural *vajzat* 'the girls' in (8d), associated with the structure in (12) (Manzini and Savoia 2011, 2018).

(12) [[[vaiz_{R}] girl a $_{\text{Class}}$] PL t $_{\text{D(Num)}}$] DEF cf. (8d)

As already mentioned, DM theorists consistently identify the lower Class node in a structure like (12) with *n*. Déchaine *et al.* (2014), working on Shona (Bantu), however take *n* to be separate from nominal class morphology, whose exponents they conceive of as Asp categories, with different flavours. I propose here that *n* is to be understood as distributed over the complex of inherent features of Ns that contribute its nominal character to the root, namely ϕ , definiteness D as well as (inherent) Case (in practice oblique case). In other words, (12) is shorthand for (13).

(13) [[[vaiz_{R}] a $_{n\text{(Class)}}$] t $_{n\text{(Def/Num)}}$] cf. (8d)

We are now in a position to consider whether linkers can be identified with *n*, which is Chomsky's (2020) proposed categorisation for the Links in adjectival sequences. Since *n* is identified with the collection of the inherent properties of N, it is evident that the Albanian linker can be construed as an exponent of *n*. Categorisations of the linker as Agr (i.e. ϕ , Philip 2012) or as Def (Lekakou and Szendrői 2012) or as D (Franco, Manzini and Savoia 2015) can be seen as approximations of the more general *n* categorisation, conceived as in (13). Formally, then, in Albanian the adjectival modifier, for instance *të bukura* 'good looking-FPL' in (8d) results from Pair-Merge of an (agreeing) adjective *bukura* and of a Link *të* of category *n*, as in (14).

(14) < [$_{n}$të], [$_{A}$bukura] > cf. (8d)
LKR.PL good-looking

The adjectival Pair-Merge structure in (14) is then merged with the N in (13) to form the A, N modification structure in (15). The structure of merger (Set-Merge and Pair-Merge) is the same as in Chomsky's English structures in section 1.[2]

(15) $[_{NP}\ [_{N}$vaizat] $<[_{n}$të], $[_{A}$bukura] >] cf. (8d)
the girls LKR.PL good looking

Another property of linkers mentioned in section 2 is that they are found not only with adjectival modification, but also with modification by an oblique case NP or by a PP. Specifically, in Albanian, linkers are present when N is modified by a genitive NP, which is the only oblique in the language. Modification of N by PPs does not involve linkers (one of many points that are left for future research). The structure of genitive modification is exactly parallel to that proposed for adjective modification, as illustrated in structure (16) for example (9c).

(16) $[_{NP}\ [_{N}$ muri] $<[_{n}$i], $[_{NP}$shtëpisë] >] cf. (9c)
the wall.M.SG LKR.M.SG of the house

Next, recall that Chomsky (2020) treats coordinations of modifiers as sequences of Pair-Merge couples, each consisting of a Link and of a modifier. Given this, we expect that it should be possible (and perhaps necessary) to find linkers repeated in front of each modifier. This is clearly the case in Albanian, as illustrated in structure (17) for example (10a) above.

(17) $[_{NP}\ [_{N}$vaiza] << e, mire >, < e, sjellshme>>] cf. (10a)
the girl LKR good LKR well-behaved

As expected, when several modifiers are present, they receive a coordinated reading. The referent of the nominal expression is at the intersection of the various predicates involved, namely 'the x: x girl and x good and x well behaved' for (17). The content of the definite operator is provided by the definite feature of the head noun. The x variable in turn is restricted by the ϕ features of *n*, in this instance, plural and (redundantly) feminine. Recall that Franco, Manzini and Savoia (2015), Manzini and Savoia (2018) propose that the linker provides a local argument for each individual predicate. Yet the reason for this ultimately eludes them. Link structures like (17) provide the required explanation, since they connect linkers to the universal structure for modification (and/or coordination), namely Pair-Merge. In this sense linker languages like Albanian do not instantiate any special construct. If anything, it is languages like English, which do not have any element externalising linkers that represent a special case.

Various empirical challenges remain open. To begin with, linkers are not simply copies of the *n* inflection of the modified N. In Albanian the linker for a modified definite N, as in (15), (16) is in fact a copy of the inflection of N. But this is not true in examples where the modified N is indefinite for instance (7a, b) or (8a, b), while the linker still belongs to the definite morphological series. In order to discuss this point, one additional fact about Albanian morphosyntax needs to be introduced, namely that Ns with so called indefinite endings can

[2] I adopt ideas by Chomsky (2013) to the effect that precedence is not represented in syntax and is computed at the externalization interface (EXT) on the basis of dominance relations in syntax. Therefore, the adjective is shown to the right of the noun in (15), or to its left in English, without prejudice to the syntactic parallelism between the two languages.

be found in combination not only with (indefinite) quantifiers but also with demonstratives. Perhaps, it would be more appropriate to speak of a weak inflection, where an overt D is present, vs. a strong inflection, conveying definiteness in the absence of a definite article. In this perspective, the strong form of linkers could reflect some morphosyntactic property, for instance their autonomous (clitic) head status. Evidently, all of this remains to be worked out. In other languages, we know that linkers are invariable (e.g. the Persian ezafe), hence possibly a default *n*. A descriptively adequate theory must ultimately take this range of variation into account.

This is also true of the constraints that restrict the overt presence of linkers in some languages. For instance, in Greek, linkers only surface when a definite N is modified. In this case, we could assume that the modification of indefinite Ns also involves linkers, except that they are not externalised. The evidence comes in part from the comparison with languages like Albanian, where linkers have a less differentiated lexicalisation when a definite N is modified. Romanian also provides an example of missing linkers dictated by externalisation and not by the absence of linkers structures. Thus, Romanian genitives are preceded by the linkers *al, a, ai, ali* which nevertheless do not surface when the modified N is definite and crucially is linearly adjacent to the modifying genitive, as shown in (18a) (no linker) vs. (18b-d) (obligatory linker, Dobrovie-Sorin, Giurgea and Nedelcu 2013).

(18) a. casa vecinului
house-F.SG.DEF neighbour-M.SG.OBL.DEF
'the neighbour's house'

b. o casă a vecinului
a house.F.SG LKR.F.SG neighbour-M.SG.OBL.DEF
'a house of the neighbour's/a neighbour's'

c. casa frumoasă a vecinului
house-F.SG.DEF beautiful LKR.F.SG neighbour-M.SG.OBL.DEF
'the beautiful house of the neighbour'

d. Casa este a vecinului
house-F.SG.DEF is LKR.F.SG neighbour-M.SG.OBL.DEF
'The house is the neighbour's.'

The final potential difficulty I consider here has to do with predicative modifiers in post copular position, preceded by the linker in Albanian (11) and in Romanian (18d). The question is how this can be reconciled with the *n* nature of the linker. The simplest answer lies in the analysis of copular constructions proposed by Moro (1997). At first Merge, a predication is created between NP and an adjective, or a possessor, or a set of modifiers. In words, the underlying structure of (11a) contains the substructure in (19) which is just a normal nominal modification structure. Merger of the copula then forces *vajza* to raise, stranding the linker structure.

(19) vajza [$_{IP}$është … [$_{NP}$ [$_{N}$ ~~vajza~~] < [$_{n}$ e], [$_{A}$bukura] >] cf. (11a)

In short, the identification of linkers in nominal modification with Chomsky's (2020) Links is consistent with empirical evidence and resolves the long-standing issue of the underlying motivation for linker structures. In the construal suggested here, linkers are an overt reflection of the ordinary structure of nominal modification, covertly present in languages like English

as well. Apart from its intrinsic interest, this conclusion opens the way for studying Pair-Merge structures on the basis of their overt linker manifestations. Several avenues of further research open up. I will briefly mention some of them in the next section.

4. Link(er)s: further prospects

Work by Collins (2019) on Khoisan languages shows the existence of verbal linkers, introducing a variety of arguments and modifiers of the verb other than the direct object. Various questions arise: first of all, whether these are instantiations of the *v* Link, and next how this complies with the coordination semantics of Link sequences. In fact, in at least one I-E, Romance language, namely Aromanian, the same linker element that precedes genitives in adnominal modification also precedes datives in ditransitive or other environments (Manzini and Savoia 2018). Interestingly, the Aromanian linker agrees with the genitive or dative, unlike all genitives seen so far.

(20) a. i o am datə o fitʃoru/ ali feti
him/her it have.1SG given LKR.M.SG boy-DEF/ LKR.F.SG girl-DEF
'I gave it to the boy/the girl'

b. libra o fitʃoru/ ali feti
book-F.SG.DEF LKR.M.SG boy-DEF/ LKR.F.SG girl-DEF
'the boy's/the girl's book'

The matter of verbal linkers is left open here. Collins (2019) himself points to the possible connection with adnominal linkers dropping the issue afterwards. The Link analysis of linkers may lead to a profitable reopening of this question.

On a different track, typological work (Plank 1995) draws together linkers and other phenomena which like linkers, characterise adnominal modification and involve the surfacing of a partial copy of the modified N on the modifier (adjective, genitive). The most notable of such phenomena is Case stacking, instantiated for instance in Australian languages. In the Lardil (Pama-Nyungan) example in (21) (Richards 2013: 43) the modified N is in the instrumental Case 'with the spear'. Its genitive modifier 'the boy's' bears not only its own Case (genitive), but also a copy of the instrumental Case of the modified N.

(21) Ngadalatha karnjin-i marun-ngan-ku maarn-ku
I spear wallaby-acc boy-gen-instr spear-instr
'I speared the wallaby with the boy's spear.'

Manzini and Savoia (2018), Manzini, Franco and Savoia (2019) argue that Case stacking responds to the same generalization that they propose for linkers, namely that the external argument of the elementary relator 'of'/Gen must have an instantiation within the relator's maximal projection, i.e. PP or (oblique) KP. As before, the issue arises why the grammar would enforce such a requirement. If the present discussion is correct, Case stacking could be conceptualised as a different realisation of the *n* linker considered in the previous section. Thus *marun-ngan-kumaarn-ku* 'with the boy's spear' would have a structure like (22) with the head *maarn-ku* 'with the spear' modified by the genitive *marun-ngan* 'the boy's' via Pair-Merge of the latter with a Link *n,* represented in this instance by a Case inflection, *ku*. This analysis however requires a more sustained discussion of Case than has been provided here, and again I leave this issue for future research.

(22) $[_{NP}<[_{KP}$marun-ngan], $[_{n}$ku]> $[_{KP}$maarn-ku]]

As a final illustration I go back to the relation of linkers to agreement, a topic raised before, but not further explored. I will begin with some examples from Italian, a language which has direct modification of Ns by adjectives, differing from English only in that adjective agrees with the N they modify, as in (23):

(23)	a.	la	bella	casa
		the.FSG	nice.FSG	home.FSG
		'the nice home'		
	b.	le	belle	case
		the.FPL	nice.FPL	home.FPL
		'the nice homes'		

A considerable amount of discussion in generative work of the last two decades has been devoted to the question whether the agreement phenomenon in (23), often referred to as Concord, is or is not to be assimilated to subject-verb agreement, hence accounted for by (a suitable version of) the minimalist rule of Agree. There are *prima facie* overwhelming reasons to assimilate Concord and verbal agreement, see Baker (2008). At the same time, Agree, beginning with Chomsky (2001) is characterised as being fundamentally asymmetrical, namely a relation between an element needing to be checked (interpreted, evaluated, deleted, etc.) and an element able to check it. If this conception is imported into Concord, a number of difficulties arise, since an example like (23) seems entirely symmetrical: every member of the NP must agree with any other member, overtly. Ways have been suggested to avoid this difficulty. In one of the first approaches to the issue, Carstens (2000) proposes checking multiple categories (A, D) by means of N. Other theorists have preferred defining separate rules (Giusti 2008 for an early proposal) or simply ignore Concord when discussing Agree.

Here I would like to add linkers and/or Links to the Agree vs. Concord equation. The Romance languages have possessive (genitive) pronouns. These pronouns of course have their own inherent ϕ features, including person and number as English *my, his, their* etc. In addition, however, they also bear an inflection agreeing with the N they modify, as shown in (24).

(24)	a.	la	mia	casa
		the.FSG	my.FSG	home.FSG
		'my home'		
	b.	le	mie	case
		the.FPL	my.FPL	home.FPL
		'my homes'		

Because of the discussion of Albanian, we know that in (23a) *bella casa* 'nice home' has the structure of embedding in (25a). We also know that genitive modifiers have a similar structure, so that *mia casa* 'my home' in (24a) corresponds roughly to (25b).

(25) a. $[_{NP}<n, [_{A}$bella] > $[_{N}$ casa]]
b. $[_{NP}<n, [_{NP}$ mia] > $[_{N}$ casa]]

The problem with (24) is the stacking of Φ features on the possessive pronoun, namely its own intrinsic features, and on top of those a copy of the features of N. Now, recall from the discussion of Albanian that the *n* linker is a copy of the *n* nominalising category of N, in practice its inflection. The representation in (25b) suggests that the stacked agreement could be another externalisation yet of the *n* Link, this time via agreement, as in (26).

(26) $[_{NP}{<}n,\ [_{NP}\ \text{mi-a}] > [_{N}\ \text{casa}]]$

What is especially interesting is that one could extend the same treatment to adjectival concord, along the lines of (27). If so, the dissimilarities between subject/verb Agree and Concord would not be a consequence of different rules applying, or of Agree applying in a different way. Rather I am suggesting that Concord may be construed as regular Agree between an *n* phase head probe and a goal, except that *n* is the Link in a modification pair merge structure.

(27) $[_{NP}{<}n,\ [_{A}\text{bell-a}] > [_{N}\ \text{casa}]]$

In fact, the connection between linkers, case stacking and what we may call agreement stacking is clearly perceived by the typological literature (Plank 1995). In the theoretical literature, Manzini, Franco and Savoia (2019) discuss in detail Punjabi where genitive PPs bear agreement with the N they modify. If (27) is correct, then a further important consequence follows from the pair merge construct involving Links (and hence linkers).

5. Conclusion

In this note, I have first presented a recent formalisation of modification and conjunction structure by Chomsky (2020) in terms of sequences of pair merge units, each unit created by means of a Link element identified with the nominaliser/verbaliser *n/v*. I have argued that the Link element is overtly visible in languages with linkers, to be more precise nominal linkers, while for verbal linkers I have left the question open. I have suggested that typologists are rights in regarding case stacking and linkers as essentially the same phenomenon. I have further suggested that concord may be another manifestation of Link structures, solving the longstanding issue of Concord vs Agree.

References

Baker, Mark C. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge UP.

Campos, Hector. 2005. "Noun Modification, Pseudo-Articles, and Last Resort Operations in Arvantovlaxika and in Romanian." *Lingua* 115 (3): 311-347.

Campos, Héctor, and Melita Stavrou. 2004. "Polydefinites in Greek and Aromanian." In *Balkan Syntax and Semantics,* ed. by Olga M. Tomić, 137-173. Amsterdam: John Benjamins Publishing Company.

Campos, Héctor. 2009. "Some Notes on Adjectival Articles in Albanian." *Lingua* 119 (7): 1009-1034. doi: 10.1016/j.lingua.2008.09.014.

Carstens, Vicky. 2000. "Concord in Minimalist Theory." *Linguistic Inquiry* 31 (2): 319-355.

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

Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework." In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, Juan Uriagereka, and Samuel Jay Keyser, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: a Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. "Problems of Projection." *Lingua* 130: 33-49.
Chomsky, Noam. 2020. "The UCLA Lectures (April 29 – May 2, 2019)." <https://ling.auf.net/lingbuzz/005485> (06/2021).
Collins, Chris. 2019. *The Linker in the Khoisan Languages*. Oxford: Oxford UP.
Déchaine, Rose-Marie, Raphaël Girard, Calisto Mudzingwa, and Martina Wiltschko. 2014. "The Internal Syntax of Shona Class Prefixes." *Language Sciences* 43: 18-46.
Dikken, Marcel D., and Pornsiri Singhapreecha. 2004. "Complex Noun Phrases and Linkers." *Syntax* 7: 1-54. doi: 10.1111/j.1368-0005.2004.00064.x.
Dikken, Marcel D. 2006. *Relators and Linkers. The Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, MA: MIT Press.
Dobrovie-Sorin, Carmen, Ion Giurgea, and Isabela Nedelcu. 2013. "Genitive DPs and Pronominal Possessors." In *A Reference Grammar of Romanian Volume 1: The Noun Phrase,* ed. by Carmen Dobrovie-Sorin, and Ion Giurgea, 309-354. Amsterdam: John Benjamins Publishing Company.
Epstein, Samuel D., Hisatsugu Kitahara, and Daniel Seely. 2016. "Phase Cancellation by External Pair-Merge of Heads." *The Linguistic Review* 33 (1): 87-102. doi: 10.1515/tlr-2015-0015.
Franco, Ludovico, M. Rita Manzini, and Leonardo M. Savoia. 2015. "Linkers and Agreement." *The Linguistic Review* 32 (2): 277-332. doi: 10.1515/tlr-2014-0024.
Giusti, Giuliana. 2008. "Agreement and Concord in Nominal Expressions." In *The Bantu-Romance Connection: A Comparative Investigation of Verbal Agreement, DPs, and Information Structure*, ed. by Cécile de Cat, and Katherine Demuth, 201-237. Amsterdam: John Benjamins Publishing Company.
Guardiano, Cristina, and Melita Stavrou. 2014. "Greek and Romance in Southern Italy: History and Contact in Nominal Structures." *L'Italia Dialettale* 75: 121-147.
Halle, Morris, and Bert Vaux. 1998. "Theoretical Aspects of Indo-European Nominal Morphology: The Nominal Declensions of Latin and Armenian." In *Mír Curad: Studies in Honor of Calvert Watkins*, ed. by Jay Jasanoff, H. Craig Melchert, and Lisi Oliver, 223-240. Innsbruck: Innsbrucker Beitraegezur Sprachwissenschaft.
Holmberg, Anders, and David Odden. 2008. "The Noun Phrase in Hawrami." In *Aspects of Iranian Linguistics*, ed. by Simin Karimi, Vida Samiian, and Donald Stilo, 129-151. Newcastle upon Tyne: Cambridge Scholars Publishing.
Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford UP.
Larson, Richard K., and Hiroko Yamakido. 2008. "Ezafe and the Deep Position of Nominal Modifiers." In *Adjectives and Adverbs. Syntax, Semantics, and Discourse*, ed. by Louise McNally, and Christopher Kennedy, 43-70. Oxford: Oxford UP.
Lekakou, Marika, and Krista Szendrői. 2012. "Polydefinites in Greek: Ellipsis, Close Apposition and Expletive Determiners." *Journal of Linguistics* 48 (1): 107-149. doi: 10.1017/S0022226711000326.
Manzini, M. Rita, and Leonardo M. Savoia. 2011. *Grammatical Categories.* Cambridge: Cambridge UP.
Manzini, M. Rita, and Ludovico Franco. 2016. "Goal and DOM Datives." *Natural Language and Linguistic Theory* 34: 197-240.
Manzini, M. Rita, and Leonardo M. Savoia. 2018. *The Morphosyntax of Albanian and Aromanian Varieties*. Berlin: de Gruyter.
Manzini, M. Rita, Ludovico Franco, and Leonardo Savoia. 2019. "Suffixaufnahme, Oblique Case and Agree." In *Agreement, Case and Locality in the Nominal and Verbal Domains*, ed. by Ludovico Franco, Mihaela Marchis Moreno, and Matthew Reeve, 211-255. Berlin: Language Science Press. doi: 10.5281/zenodo.3458074.
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: 201-25.
Moro, Andrea. 1997. *The Raising of Predicates*. Cambridge: Cambridge UP.
Philip, Joy N. 2012. *Subordinating and Coordinating Linkers*. PhD thesis. London: UCL.
Plank, Frans. 1995. "(Re-)Introducing Suffixaufnahme." In *Double Case. Agreement by Suffixaufnahme*, ed. by Frans Plank, 3-112. Oxford: Oxford UP.

Richards, Norvin. 2010. *Uttering Trees*. Cambridge, MA: MIT Press.
Richards, Norvin. 2013. "Lardil 'Case Stacking' and the Timing of Case Assignment." *Syntax* 16 (1): 42-76. doi: 10.1111/j.1467-9612.2012.00169.x.
Turano, Giuseppina, 2002. "On Modifiers Preceded by the Article in Albanian DPs." *University of Venice Working Papers in Linguistics* 12: 169-215.
Turano, Giuseppina. 2004. *Introduzione alla grammatica dell'albanese*. Firenze: Alinea Editrice.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

OPEN ACCESS

Citation: G. Samo (2021) N-merge systems in adult and child grammars: a quantitative study on external arguments. *Qulso* 7: pp. 103-130. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12005

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

N-merge systems in adult and child grammars: a quantitative study on external arguments

Giuseppe Samo
Beijing Language and Culture University (<samo@blcu.edu.cn>)

Abstract:

This paper explores quantitative results based on theoretical assumptions related to the predictions on N-merge systems (Rizzi 2016) ranked from minimum to a maximum of complexity in terms of the computational devices and derivational operations they require. We investigate the nature of external arguments focussing on 2-merge systems (two elements of the lexicon merge and the created unit is again merged with a further element directly extracted from the lexicon) and 3-merge systems (merge two elements created by previous operations of merge). We add a quantitative dimension to the established qualitative dimension discussed in the theory (Rizzi 2016) by investigating large-scale corpora representative of three populations of speakers: adult grammar (102 treebanks/101 languages), typically developing children (2 corpora/English and Chinese) and children with atypical development (1 corpus). The results confirm the predictions in Rizzi (2016): every language in our data set exploits 3-merge systems and less complex systems are the preferred options in early grammars.

Keywords: *Large Datasets, N Merge System, Quantitative Syntax, Subject*

1. Introduction

This paper explores quantitative results based on theoretical assumptions related to the predictions of N-merge systems (Rizzi 2016: 142), systems ranked from minimum to a maximum of complexity in terms of the computational devices and derivational operations they require, in adult grammar and two developmental populations based on large-scale corpora (discussed in details in sub-section 2.2). In particular, we here develop a series of (theory-internal) queries to extract frequencies of 2-merge systems (two elements of the lexicon merge and the created unit is again merged with a further element directly extracted from the lexicon) and 3-merge systems (merge between two elements created by previous operations of merge) in large-scale corpora.

We focus our search on a specific syntactic phenomenon, namely the nature of external arguments (henceforth, EA). EA are assumed to be generated in the verbal domain (vP) of transitive and unergative verbs (Perlmutter 1978; Belletti and Rizzi 1981; Chomsky 1981; Hale and Keyser 1998; for a discussion on externality of arguments; Gallego 2008; Preminger 2008; for more fine-grained cartographic analyses of the vP layer, Ramchand 2008; Si 2019). If no further operation is involved (e.g., smuggling in the case of passivization, Collins 2005; Belletti and Collins 2021), the functional projection hosting EAs in the vP of transitive and unergative verbs are to be analysed as the locus of generation of the "classical" notion of clausal subjects (which will further move to a dedicated criterial position, see Rizzi 2015b: 17ff, for an overview) in these architectures.

EAs represent a first point of investigation for quantitative analyses of N-merge systems since basic layers of complexity arise. Let us compare, for example, the English sentences in (1a) and (1b), where the verbal element *write-* stands for an uninflected form generated within the vP layer.

(1) a. $[_{\text{HEAD}}$ She $[_{\text{XP}}$ $[_{\text{HEAD}}$ write-] $[_{\text{XP}}$ the paper]]]
b. $[_{\text{XP}}$ [The young linguist] $[_{\text{XP}}$ $[_{\text{HEAD}}$ write- $[_{\text{XP}}$ the paper]]]]

In (1a), a pronominal element *she* merges with an already formed complex phrase composed of *write-* and *the paper*, while in (1b) the EA of the verb is a complex element created by the combination of the definite article *the*, the adjective *young* and the noun *linguist*, which then need to be merged with another complex element *write the paper*. These two configurations bear different layers of difficulty (Rizzi 2015a, 2015b, 2016), as will be discussed throughout the work, and represent two types of N-merge systems: (1a) is a case of 2-merge, while (1b) is a case of 3-merge.

In this study, we provide a quantitative dimension to the established qualitative dimension discussed in Rizzi (2016), investigating large-scale corpora representative of three populations of speakers: adult grammar corpora (102 syntactically annotated treebanks for 101 languages, Universal Dependencies, Nivre 2015; Zeman *et al.* 2020), typically developing (TD) children and children with atypical development datasets (4 morpho-syntactically annotated corpora in Childes; MacWhinney 2000a, 2000b) in order to determine if asymmetries exist between these populations (cf. Borer and Wexler 1987; Guasti 2002, 2017; Kam and Newport 2005; Friedmann, Belletti and Rizzi 2009; Durlemman *et al.* 2015; Stanford 2020, ch.1 for an overview).

We aim to detect if occurrences of generation of complex external arguments (3-merge system, in 1b) can be found in all the adult grammar corpora under investigation and if asymmetries in distributions between 3-merge (1b) and 2-merge (1a) systems arise, in terms of frequencies, in both adult and child grammar. Following Merlo and Stevenson (1998: 134), frequency is intended here as a measure to quantify a role played by grammar (Merlo and Stevenson 1998; Bresnan, Dingare and Manning 2001; Merlo 2016; Samo and Merlo 2019; Gulordava and Merlo 2020). We aim to show that the theoretical predictions on the status of N-merge (Rizzi 2016) quantitatively result in frequencies. This work should represent a first step towards a more complex mapping between theoretical assumption on the complexity of structures and observational data extracted from large datasets.

In section 2, we present a typology of N-merge computational systems (based on Rizzi 2016) which can be naturally ranked from minimum to a maximum of complexity in terms of the computational devices and derivational operations they require. Section 3 presents relevant methods in quantifying complexity with respect to observed frequencies: we verify the

consistency of the predictions of the typology postulated in Rizzi (2016) with the data available from corpus studies in large-scale on a wide range of natural languages syntactically-annotated treebanks and selected corpora from language acquisition. In section 4, we then turn to empirical evidence extracted from syntactically annotated adult treebanks. Once baseline conditions have been established in adult grammar, we investigate two child grammar corpora in English and Chinese, and an English corpus of children with atypical development in section 5. Finally, section 6 discusses and concludes.

2. External arguments and N-merge systems

2.1 Some notes on the cartography of external arguments and subject positions

A wealth of literature has investigated the formal nature of subject and its comparative dimensions (Rizzi 1982; Jaeggli and Safir 1989; Biberauer *et al.* 2010 among others). The subject is considered as an obligatory element (Chomsky 1981) generated as an argument inside the vP (Larson 1988; Borer 1994; Chomsky 1995; Si 2019), which then moves to the Specifier of the Inflectional layer IP/Tense layer TP (Pollock 1989; Koopman and Sportiche 1991; Cardinaletti 2004; Rizzi 2006, 2015a, 2015b; Berthelot 2017) to satisfy subject requirements (e.g., subject criterion in Rizzi 2006; 2015b as a reformulation of classical Chomsky 1981's Extended Projection Principle).

In languages like English, the EA undergoes obligatory movement to a dedicated criterial position SubjP (Cardinaletti 2004; Rizzi 2006) shown in (2).[1] This contrasts with the lack of obligatory movement of the other argument generated in transitive constructions, the 'object of the verb', referred to as the 'internal argument' (IA).

(2) [SpecSubjP The linguist [EA <The linguist> [vP write- [IA the paper]]]]

In transitive verb constructions, if passivization is not involved (in terms of smuggling, Collins 2005) the EA represents the element which will undergo movement to the dedicated functional projection.

Naturally, subjects target other positions if they bear the relevant features. For example, subjects can enter the syntax of cleft structures (Belletti 2015) and relative clauses (Cinque 2014). If the elements bear relevant features (e.g., +Top, +Focus, +Q), subjects can move to dedicated functional projections for topic and focus positions (Frascarelli and Hinterhölzl 2007; Bianchi and Frascarelli 2010; Bianchi, Bocci and Cruschina 2015; Bonan 2019) within both the Left Periphery (Rizzi 1997; Rizzi and Bocci 2017) and the low vP-periphery (Belletti 2004) of the clause. The locus of generation and the landing site of the movement is crucial for explaining asymmetries of subject vs. non-subject configurations in adult processing (Frauenfelder, Segui and Mehler 1980; Chesi and Canal 2019), in TD children (Friedmann, Belletti and Rizzi 2009; Belletti *et al.* 2012), in atypical development (Durrleman *et al.* 2015; Stanford 2020) and in language pathology (e.g., aphasic patients, Grillo 2008; Martini *et al.* 2020; Alzheimer's Disease, Caloi 2013).

[1] A finer cartography of subject positions for different subject elements is provided in Cardinaletti (2004: 116, 154-156).

Linguistic variability can be also detected by the realization of the overt realization of the subject position. Among the syntactic strategies available, languages may drop the subject in certain contexts (Rizzi 1982; Haegeman 1990; Frascarelli 2007; for a typology of languages see Biberauer *et al.* 2010). This will be further discussed in section 4.

2.2 A typology of N-merge systems

Rizzi (2016), rediscussing results of decades of study in generative grammar, takes as starting point the idea that a language system is thought to require at least three main components. The first component is the lexicon, formed by a finite list of items representing linguistic items. The second component is a combinatorial device merge, whose role is to create complex expressions with elements extracted from the lexicon. The operation of "merge" (Chomsky 1957, 1995, 2013; Moro 1997; discussed as a biological primitive in Dehaene *et al.* 2015; see also Boeckx and Theofanopoulou 2014; Murphy 2015), forms a minimal binary tree by combining two elements in order to create a new linguistic object. Finally, a hypothetical (set of) working space(s) is required to apply reiterated operations of merge. Given such components, Rizzi identifies a typology of merge systems with a hierarchy of complexity, "ranked in terms of their generative capacity and of the computational resources they need" (Rizzi 2016: 142).

We here introduce the notion of head [head] and the notion of maximal projection [XP]. Following Rizzi (2015b: 20), we take heads to be directly extracted from the lexicon and maximal projections to be syntactic objects created by merge operations.[2]

Rizzi (2016: 143) proposes a transparent typology of systems labelled according to the number of merge operations: 0-merge, 1-merge, 2-merge and 3-merge systems, with the latter two as objects of our investigation.

In 0-merge systems, an element called head is directly extracted from the lexicon and sent to the systems of sounds and meaning. Only one device is exploited and there are no instances of merge. This system predicts languages based on single word utterances, where head means directly extracted from the lexicon. This system could represent the basis of a subset of linguistic systems, such as those, following Rizzi (2016: 144), belonging to a large set of monkey populations (see Schlenker *et al.* 2016 for an in-depth study) and of early phases of language acquisition (one-word utterances, following Guasti 2002: 24 and citation therein).

1-merge is a system that uses both devices: two elements of the lexicon are merged creating a phrase. Once the phrase is built, it is sent directly to the interface without any recursive procedures. This system could represent an option in a subset of other animals' linguistic systems, such as, following Rizzi (2016 on comment of Schlenker *et al.* 2016), the morphosyntax of a smaller set of monkey populations which produce some rudimentary forms of combinatorial systems.

A 2-merge system is a system in which two elements taken from the lexicon are merged, and a second working space combines a third element from the lexicon. This system is more complex than 0-merge and 1-merge systems, but still relatively simple. Two elements of the lexicon merge and the created unit is again merged with a further element directly extracted from the lexicon. Even if the system seems to be a perfect match between economy of computation and complexity, it does not capture the behaviour of natural languages. Moreover, this system would only lead to either uniformly right or uniformly left tree structures. As expected, such a conclusion might prove to be descriptively inadequate, as predicted by Rizzi (2016: 143).

[2] The discussion on branching direction (left- vs. right- branched languages, see Kayne 1994) and the direction of phrase structure building (bottom-up vs. top-down, see Chesi 2012, 2015) is beyond the scope of this paper.

Finally, 3-merge systems are able to merge two elements created by previous operations of merge. This requires two working systems, lexicon and the operation of merge.

Following the description provided by Rizzi "human languages manifest the full power of 3-merge systems [XP, XP]: no human language is limited to the use of single words (0-merge) or just two-word sequences (1-merge, [head-head]), or to disallow complex specifiers (2-merge, [head, XP]). In other words, human systems possess all the three configurations" (Rizzi 2016: 144): this "human" nature discussed in Rizzi (2016) restricts our field of interest to the last two systems only, 2-merge and 3-merge.

3. Quantifying Hypotheses

A preliminary research question of this paper is directly extracted by Rizzi (2016: 144).

> A 2-merge system would only permit external arguments consisting of one word like *[he [will [meet [the girl]]]]*, but not of two words like *[[the boy] [will [meet [the girl]]]]* (a structure which would require the power of a 3-merge system): no human language appears to have this limitation and disallow complex specifiers. (2016: 144)

The study addresses the prediction in large-scale corpora to investigate whether this conclusion can be made. Following Merlo (2016 and related works), we use corpus counts in the spirit of the computational quantitative syntax framework, observing differentials in frequencies as the expression of underlying grammatical properties. Frequency is here adopted as a measure to quantify linguistic proposals (on the role of frequency in grammar, acquisition and formal rules; Yang 2013, 2015; see Ibbotson 2013 for a usage-based grammar account) and represents a dependent variable to test linguistic models (Merlo 2016 and related works).

The first research question then investigates a linguistic evidence from a large set of languages, expecting that in no language is the 3-merge system absent. In other words, every language should have at least one occurrence of an external argument built with a 3-merge system. We state the research question in H_1.

H_1: The frequency of 3-merge systems in adult grammar corpora should be $\neq 0$ in every language under investigation.

Rizzi (2016) does not predict any preference, but the two systems should be equally exploited by adult grammars. In this paper, we contribute with a quantitative analysis of these two elements to observe whether we can observe a preference for 3-merge systems over 2-merge systems or vice versa.

Our second research question involves a dimension in terms of development in acquisition. Data from corpora and experimental studies in language acquisition clearly show that 2- and 3-merge systems co-exist at early stages of child grammar (Guasti 2002, 2017 and Belletti and Guasti 2015, for an overview; Friedmann, Belletti and Rizzi 2020). However, the exploitation of a different number of computational devices (2-merge: lexicon, merge, one working space; 3-merge: lexicon, merge, two working spaces) results in asymmetries. Along the same lines, we expect that, if there is any (non-marginal) increase in the production, it should affect 2-merge systems earlier than 3 merge systems.

H_2: The distributions of emergence of 2- and 3-merge systems in child grammar should differ.

Finally, we also investigate a corpus of atypical development to observe patterns whether similar asymmetries can be found.

In order to quantitatively answer these questions, we extract data from large-scale resources. To facilitate our search, we rely on large (morpho-)syntactically annotated corpora. The materials and the methods in retrieving the frequencies of the two systems are presented in the relevant sections and sub-sections.

Following Merlo (2016) and related works, our quantitative hypotheses presented here are to be contrasted to a H_0 hypothesis that would predict that grammatical properties are uncorrelated to frequencies.

To answer H_1, we investigate corpus evidence from adult grammar presented in section 4. As for H_2, we extract our data from child spontaneous production repositories presented in section 5.

4. A crosslinguistic study in Adult Grammar

The size of linguistic materials on adult grammar are huge and heterogeneous, therefore we establish a set of parameters in choosing the appropriate material for our research. First of all, our goal is to automatically gather as much linguistic evidence as possible. In order to fully automatize our search and make it replicable at different layers, we chose (morpho)-syntactically annotated corpora. A fundamental annotation is in terms of the grammatical (syntactic) function in the sentence, to detect occurrences of external arguments (subjects in most syntactic annotations' tools). If the grammatical functions are possibly combined with a Part-of-Speech tag (henceforth, PoS) annotation, it is faster to classify external arguments (subjects) as maximal projections (noun, proper nouns) and heads (pronominal entities). Merging these two elements provides the right amount of information to detect and differentiate the two merge systems under investigation. Finally, we used a syntactically annotated database allowing us to investigate as many languages (and language families) as possible. Candidates are large, multilingual, homogeneously-annotated data sets, as the treebanks provided by the Universal Dependencies (Nivre 2015; Zeman *et al.* 2020), which will be briefly described in sub-section 4.1.

4.1 Materials and Methods

Our material is extracted from syntactically annotated treebanks following the guidelines of the Universal Dependencies (henceforth UD, version 2.7, Zeman *et al.* 2020) annotation scheme, allowing direct comparison across languages. We take into consideration 102 treebanks for 101 languages. To factor out problems related to genre classification of the treebanks, we chose, when possible, the biggest and the most heterogeneous treebank for each language. A parameter in preferring a specific treebank was given by the number and the quality of types of registers provided by UD guideline screen.[3] We avoided treebanks with less than 50 trees and, when possible, parallel treebanks (Ahrenberg 2007; Volk, Graën and Callegaro 2015). Detailed information of the materials (size, references) is provided in Table 1 together with the results of study 1.

[3] <http://www.universaldependencies.org > (06/2021).

For Italian, we adopted two treebanks to detect, if any, genre effects on the distribution of complex structures (in the spirit of Samo, Zhao, and Gamhewage 2020 and Zhao *et al.* 2021): the ISDT treebank v.2.7 (Bosco, Montemagni and Simi 2013, text genres: legal, news, wiki) and TWITTIRO treebank 2.7 (Cignarella *et al.* 2018, text genres: social media).

Beyond investigating the presence of 3-merge systems in 97 monolingual treebanks, we also controlled for specific populations: a sign language treebank (Swedish Sign Language), two bilingual treebanks (Hindi-English and Turkish-German), and two treebanks of learner essays of L2 (English and Chinese L2). For recent generative (cartographic) analysis on these populations, see Bross (2020) for Sign Languages; Shim (2016) for code-switching; and Di Domenico, Baroncini and Capotorti (2020) for L2, and references therein.

We performed two studies. In the first study, we observed the distribution of 3-merge systems only. In the second study, we aimed to quantify whether there is a preference between the two systems. In this latter study, we also took into consideration the role of null subject, following typological classifications provided by the World Atlas of Language Structures (WALS, Dryer and Haspelmath 2013), to evaluate our data in light of the availability of choices in a given language.[4] We considered non-null subject languages those labelled as "Obligatory pronouns in subject position" in Dryer (2013) and WALS as null subject languages, whereas the set of labels ("Subject clitics on variable host", "Subject affixes on verb", "Optional pronouns in subject position" and "Mixed") as languages allowing null subject constructions, but we do not discuss the different fine-grained differences among the different groups (see, for example, Neeleman and Szendrői 2007; Biberauer *et al.* 2010; Holmberg and Roberts 2013; Frascarelli and Casentini 2019). We analysed 48 treebanks for 47 languages, since both treebanks of Italian are investigated to detect, if any, further effects of text genre.

Finally, to detect cross-linguistic genre effects, we ran a linear regression (inspired by Merlo and Ouwayda 2018) and automatically ranked the costs of specific genres in the distribution of the treebank. We used the Waikato Environment for Knowledge Analysis, WEKA v.3.8.2 (Hall *et al.* 2009) to derive the best linear regression model of this data. Each treebank is encoded as a vectorial representation and every genre is encoded as an indicator variable: 0 and 1 indicates if the genre is present or not in the relevant treebank. Positive and negative coefficients indicate the difference from a predicted frequency: genre that are associated with lower distributions will have negative coefficients, and those associated with higher distributions will get positive coefficients. We used a leave-one-out cross-validation.

All the data were extracted with the online web-based tool Match Grew.[5] We followed the guidelines for mapping UD into cartographic representation (and viceversa) proposed in Samo (2019).

Among the different syntactic labels, provided by the annotation scheme, we focused only on the core dependency subject, represented as *nsubj*: as for the modifiers of the subject, we looked at the dependencies starting from the subjects provided by *det* 'determiner', *case* 'case marking' and *amod* representing any nominal modifiers. Another important ingredient is the annotation of the PoS tag: we were therefore able to detect the pronominal or full nominal nature of the external argument.

The query providing 3-merge occurrences provided the occurrence of at least one complex specifier (modifiers) of subjects (3-merge systems) in bi-argumental/transitive constructions (the verb governs a subject and an object dependency). The query detected all of the occurrences

[4] <https://wals.info/feature/101A#2/18.0/148.2> (06/2021).

[5] <www.grew-match.fr> (06/2021)

of sentences given a variable being dependent of the dependency *nsubj* (subject) and governor of one of the set of dependencies *det* (determiner), *case* (case marking), *amod* (modifier such as adjectives). To this count, we added the frequencies of subjects annotated as PROPN (proper nouns) as PoS in transitive constructions. Proper nouns have been considered XP, following theoretical considerations in Longobardi (1994: 641 and related reference), since the nominal element is described as targeting a D position within the DP. Naturally, the utterances retrieved may have a higher layer of complexity, involving the presence of other arguments (e.g., indirect objects), adverbials and complements.

For 2-merge, we developed a query which retrieved dependents on a subject dependency, whose part-of-speech is a pronoun or bare nominals and which do not govern any modification (*det, case, amod*) dependency.

The queries for the adopted tool and a naturally occurring sentence for type extracted from Italian (ISDT treebank, Bosco, Montemagni and Simi 2013), are given in (3).[6]

(3) Queries (grewmatch.fr, accessed, 05.06.2021) and naturally occurring examples in Italian

a. 3-merge systems for subjects

i. pattern {a -[nsubj]-> b; b -[det|case|amod]-> c; b [upos = "NOUN"]; a-[obj]->d}

ex. i pompieri hanno isolato la sala. (ISDT-isst-tanl-7)
the firefighters have seal.off the room
'Firefighters sealed off the room'

ii. pattern {a -[nsubj]-> b; b [upos = "PROPN"]; a-[obj]->c}

ex. Moretti rappresenta il cinema di oggi (ISDT-isst-tanl-1511)
Moretti represents the cinema of today
'Moretti represents today's cinema'

b. 2-merge systems for subjects

i. pattern {a -[nsubj]-> b; b [upos = "PRON"]; a-[obj]->d} without {b -[det|case|amod]-> c}

ex. Io studio l' inglese (ISDT-isst-tanl-2961)
I study the English
'I study English'

ii. pattern {a -[nsubj]-> b; b [upos = "NOUN"]; a-[obj]->d } without {b -[det|case|amod]-> c }

ex. Graffiti imbrattano le città. (ISDT-isst-tanl-1924)
Graffiti litter the cities
'Graffiti litter cities'

Results are discussed in sub-section 4.3.

4.3 Results

Table 1 summarizes the results of study 1 confirming H1 (The frequency of 3-merge systems in adult grammar corpora should be ≠ 0 in every language under investigation). As Table 1 shows, every language under investigation, even in smaller treebanks, exploits 3-merge

[6] The tool adopted in the investigation (accessed February 18th, 2021) provided only the first 1000 occurrences of the query, a coefficient has been calculated on the basis of the occurrences. This coefficient is calculated to provide a better understanding of a predictive tool. The trees are used as coefficients instead of subjects to keep an analysis in terms. Being *I* an imputed count, *F* the frequency of the result, and *C* the percentage of the exploitation of the corpus, the imputed count is derived from the formula. $I = F / C$.

systems (see raw frequencies in column FREQ(UENCY) 3-MERGE and the distribution among trees in column FREQ(UENCY)/TREES). Table 1 shows the results for 102 treebanks, confirming H_1.

LANGUAGE	TREEBANKS GENRE	TREES	TOKENS	FREQ. 3-MERGE	FREQ/ TREES	References (if website, https://universaldependencies.org/treebanks/[...])
Afrikaans	AfriBooms L, NF	1934	51210	445	0.23	[…]/af_afribooms/index.html
Akkadian	RIAO NF	1799	23701	132	0.07	Luukko *et al.* (2020)
Albanian	TSA W	60	982	7	0.12	[…]/sq_tsa/index.html
Amharic	ATT B, F, GE, N, NF	1074	11084	121	0.11	Ephrem Seyoum, Miyao and Yimam (2018)
Ancient Greek	PROIEL B, NF	17080	231079	1937	0.11	Haug and Jøhndal. (2008)
Apurina	UFPA N, NF	75	635	7	0.09	Freitas (2017)
Arabic	NYUAD N	19738	758627	12466	0.63	[…]/ar_nyuad/index.html
Mod. East. Armenian	ArmTDP BL, F, GE, L, N, NF	2502	55132	244	0.10	Yavrumyan (2019)
Mod. Stand. Assyrian	AS N, NF	57	510	3	0.05	[…]/aii_as/index.html
Bambara	CRB N, NF	1026	14849	34	0.03	[…]/bm_crb/index.html
Basque	BDT N	8993	130436	1153	0.13	[…]/eu_bdt/index.html
Belarusan	HSE F, L, N, NF, P, SM	23534	298867	1283	0.05	[…]/be_hse/index.html
Bhojpuri	BHTB N, NF	357	7022	29	0.08	Ojha and Zeman (2020)
Breton	KEB F, GE, N, NF, P, W	888	10942	88	0.10	Tyers and Ravishankar (2018)
Bulgarian	BTB F, L, N	11138	167287	1327	0.12	[…]/bg_btb/index.html
Buryat	BDT F, GE, N	927	11112	49	0.05	Badmaeva and Tyers (2017)
Cantonese	HK S	1004	14922	31	0.03	[…]/yue_hk/index.html
Catalan	AnCora N	16678	548649	18088	1.08	[…]/ca_ancora/index.html

Chinese	GSD [W]	4997	128288	954	0.19	[...]/zh_gsdsimp/index.html
Chinese - L2	CFL [LE]	451	7707	26	0.06	[…]/zh_cfl/index.html
Chukchi	HSE [S]	1004	7211	14	0.01	[...]/ckt_hse/index.htm[l]
Classical Chinese	Kyoto [NF]	48434	281556	3990	0.08	[...]/lzh_kyoto/index.html
Sahidic Coptic	Scriptorium[B, F, NF]	1873	50504	153	0.08	[...]/cop_scriptorium/index.html
Croatian	SET [N, WEB, W]	9010	208419	2222	0.25	Agić and Ljubešić (2015)
Czech	PDT [N, NF, R]	87913	1596965	12808	0.15	Bejček *et al.* (2014)
Danish	DDT [F, N, NF, S]	5512	106245	954	0.17	Johannsen, Alonso and Plank (2015)
Dutch	Alpino [N]	13578	222179	3146	0.23	[…]/nl_alpino/index.html
English	EWT [BL, EM, R, SM]	16622	271478	1681	0.10	[…]/en_ewt/index.html
English-L2	ESL [LE]	5124	102805	339	0.07	[…]/en_esl/index.html
Erzya	JR [F]	1690	18838	107	0.06	Rueter and Tyers (2017)
Estonian	EDT [A, F, N, NF]	30972	469143	2441	0.08	[…]/et_edt/index.html
Faroese	OFT [W]	1208	11210	24	0.02	Tyers *et al.* (2018)
Finnish	FTB [GE]	18723	178335	952	0.05	[…]/fi_ftb/index.html
French	GSD [BL, N, R, W]	16341	416740	3865	0.24	[…]/fr_gsd/index.htm[l]
Galician	CTG [L, M, N, NF]	3993	142830	2562	0.64	[…]/gl_ctg/index.html
German	HDT [N, NF, R, WEB, W]	189928	3589318	85216	0.45	Borges Völker *et al.* (2019)
Gothic	PROIEL [B]	5401	60737	240	0.04	[…]/got_proiel/index.htm[l]
Greek	GDT [N, S, W]	2521	65962	931	0.37	Prokopidis and Papageorgiou (2017)
Hebrew	HTB [N]	6216	167633	522	0.08	Tsarfaty (2013)
Hindi	HDTB [N]	16647	368351	5930	0.36	Bhat *et al.* (2017)

Hindi-English	HIENCS [SM]	1898	28807	185	0.10	Bhat *et al.* (2018)
Hungarian	Szeged [N]	1800	43832	517	0.29	[…]/hu_szeged/index.htm[l]
Icelandic	PUD [N, W]	1000	19833	171	0.17	[…]/is_pud/index.htm[l]
Indonesian	GSD [BL, N]	5593	127516	1451	0.26	[…]/id_gsd/index.html
Irish	IDT [F, L, N, W]	4910	120879	1044	0.21	[…]/ga_idt/index.html
Italian	ISDT [L, N, W]	14167	312547	3166	0.22	Bosco, Montemagni and Simi (2013)
Italian Twitter	TWITTIRO [SM]	1424	31029	295	0.21	Cignarella *et al.* (2018)
Japanese	GSD [BL, N]	8071	200676	1135	0.14	[…]/ja_gsd/index.html
Karelian	KKPP [N, NF, WEB]	228	3322	10	0.04	[…]/krl_kkpp/index.html
Kazakh	KTB [F, N, W]	1078	11614	63	0.06	Makazhanov *et al.* (2015)
Komi Permyak	UH [F]	81	920	2	0.02	Reuter, Partanen and Ponomareva (2020)
Komi Zyrian	Lattice [F]	435	5437	14	0.03	Partanen *et al.* (2020)
Korean	Kaist [A, F, N]	27363	377453	523	0.02	[…]/ko_kaist/index.html
Kurmanji	MG [F, N]	754	11014	39	0.05	Gökırmak and Tyers (2017)
Medieval Latin	ITTB [NF]	26977	477492	1731	0.06	Cecchini *et al.* (2018)
Latin	PROIEL [B, NF]	18411	218574	877	0.05	[…]/la_proiel/index.html
Latvian	LVTB [A, F, L, N, S]	13643	234179	1186	0.09	[…]/lv_lvtb/index.html
Lithuanian	ALKSNIS [F, L, N, NF]	3642	73693	153	0.04	Bielinskienė *et al.* (2016)
Livvi	KKPP [N, NF, WEB]	125	1757	6	0.05	[…]/olo_kkpp/index.htm[l]
Maltese	MUDT [F, L, N, NF, W]	2074	46236	352	0.17	[…]/mt_mudt/index.html
Manx Gaelic	Cadhan [B, BL, F, N, NF, SM, WEB, W]	291	6445	17	0.06	[…]/gv_cadhan/index.htm[l]

Marathi	UFAL [F, W]	466	4315	16	0.03	[...]/mr_ufal/index.htm[l]
Mbya Guarani	Thomas [NF]	98	1416	2	0.02	[...]/gun_thomas/index.htm[l]
Moksha	JR [N, NF]	167	1681	3	0.02	Reuter (2018)
Munduruku	TuDeT [N, NF]	62	333	1	0.02	[...]/myu_tudet/index.ht[ml]
Naija	NSC [S]	9242	149971	502	0.05	[...]/pcm_nsc/index.html
North Sami	Giella [N, NF]	3122	29967	151	0.05	Tyers and Sheyanova (2017)
Norwegian Bokmål	Bokmaal [BL, N, NF]	20044	330265	1584	0.08	[...]/no_bokmaal/index.htm[l]
Norwegian Nynorsk	Nynorsk [BL, N, NF]	17575	318928	2342	0.13	Velldal, Øvrelid and Hohle (2017)
Old Church Slavonic	PROIEL [B]	6338	63901	231	0.04	[...]/cu_proiel/index.html
Old French	SRCMF [L, NF, P]	17678	188418	1914	0.11	Stein and Prévost (2013)
Old Russian	TOROT [L, NF]	16944	166724	827	0.05	Eckhoff and Berdičevskis (2015)
Persian	Seraji[F, L, M, N, NF, S, SM]	5997	158917	435	0.07	[...]/fa_seraji/index.html
Polish	PDB [F, N, NF]	22152	372188	1539	0.07	[...]/pl_pdb/index.html
Portoguese	GSD [BL, N]	12078	331931	3938	0.33	[...]/pt_gsd/index.html
Romanian	RRT [A, F, L, M, N, NF, W]	9524	228035	1138	0.12	[...]/ro_rrt/index.html
Russian	SynTagRus___[F, N, NF]	61889	1169630	4990	0.08	Droganova, Lyashevskaya and Zeman(2018)
Sanskrit	Vedic [NF]	3997	31114	71	0.02	Hellwig *et al.* (2020)
Scottish Gaelic	ARCOSG [F, N, NF, S]	3173	63590	131	0.04	Batchelor (2019)
Serbian	SET [N]	4384	102057	1045	0.24	[...]/sr_set/index.html
Skolt Sami	Giellagas [N, NF, S]	104	1456	4	0.04	[...]/sms_giellagas/index.html
Slovak	SNK [F, N, NF]	10604	116734	1072	0.10	Zeman (2017)
Slovenian	SSJ [F, N, NF]	8000	148670	1253	0.16	[...]/sl_ssj/index.html

Spanish	AnCora [N]	17680	567429	11633	0.66	[...]/es_ancora/index.html
Swedish	Talbanken [N, NF]	6026	102884	706	0.12	[...]/sv_talbanken/index.html
Swedish Sign	SSLC [S]	203	1813	6	0.03	[...]/swl_sslc/index.html
Swiss German	UZH [BL, F, N, NF, W]	100	1544	17	0.17	Aepli and Clematide (2018)
Tagalog	Ugnayan [F, NF]	94	1191	19	0.20	[...]/tl_ugnayan/index.html
Tamil	TTB [N]_	600	10181	73	0.12	[...]/ta_ttb/index.html
Telugu	MTG [GE]	1328	7739	57	0.04	[...]/te_mtg/index.html
Thai	PUD [N, W]	1000	23322	139	0.14	[...]/th_pud/index.html
Turkish	BOUN [N, NF]	9761	132144	649	0.07	Türk *et al.* (2020)
Turkish-German	SAGT [S]	1891	33837	51	0.03	Çetinoğlu and Çöltekin (2019)
Ukrainian	IU [BL, EM, F, GE, L, R, SM, WEB, W]	7060	129384	713	0.10	[...]/uk_iu/index.html
Upper Sorbian	UFAL [NF, W]	646	11842	94	0.15	[...]/hsb_ufal/index.html
Urdu	UDTB [N]	5130	143207	2199	0.43	Bhat *et al.* (2017)
Uyghur	UDT [F]	3456	43962	160	0.05	[...]/ug_udt/index.htm[l]
Vietnamese	VTB [N]	3000	46754	514	0.17	[...]/vi_vtb/index.htm[l]
Warlpiri	UFAL [GE]	55	369	4	0.07	[...]/wbp_ufal/index.html
Welsh	CCG [F, GE, N, NF, W]	1657	34568	34	0.02	Heinecke and Tyers (2019)
Wolof	WTB [B, W]	148	46365	319	2.16	[...]/wo_wtb/index.html
Yoruba	YTB [B, W]	318	8561	67	0.21	[...]/yo_ytb/index.html

Table 1 – Treebank, Size (trees and tokens), raw frequencies of 3-merge and distribution of 3-merge out of number of trees in every language under investigation. Genre abbreviations: A = Academic, B = Bible, BL = Blog, EM = Emails, F = Fiction, G = Government, GE = Grammar Examples, L = Legal, LE = Learner-essays, M = Medical, N = News, NF = Nonfiction, P = Poetry, S = Spoken, SM = Social Media, R = Reviews, W= Wikipedia, WEB = Web

The distribution of 3-merge with respect to the trees on the treebank does not correlate with the size of the treebanks (r = 0.18). As for text genre in Italian, the results show minimal differences between the distribution of 3-merge systems in the two investigated treebanks (ISDT 0.22, Twittiro 0.21). The Linear Regression model's results (r = 0.03; Academic +0.3268, Web +0.1076, Fiction -0.146), shows a marginal increase of 3-merge in Academic and Web extracted material, whereas fiction (e.g., novels) seems to have a negative impact on the production of 3-merge systems. Future studies should investigate dimensions of variation among genres.

The second study investigated whether there is a preference between the two systems (2-merge and 3-merge). Preference is intended here in frequencies. Table 2 summarizes the results.[7]

Language	Null (p)	Freq3m	%3m	Freq2m	%2m	Pref. (k)	*Obj (N)*	*z-test*
Albanian	Yes	7	0.37	12	0.63	2merge	*34*	p = .459329.
Amharic	Yes	121	0.24	377	0.76	2merge	572	p < .000001.
Arabic	Yes	12466	0.51	11998	0.49	3merge	118666	p < .000001.
Armenian	Yes	244	0.33	506	0.67	2merge	2338	p < .000001.
Basque	Yes	1153	0.51	1114	0.49	3merge	7522	p < .000001.
Breton	Yes	88	0.77	27	0.23	3merge	393	p = .000005.
Bulgarian	Yes	1327	0.40	2000	0.60	2merge	6271	p = .032073.
Catalan	Yes	18088	0.76	5651	0.24	3merge	37089	p < .000001.
Chinese	Yes	954	0.33	1934	0.67	2merge	7748	p < .000001.
Croatian	Yes	2221.5	0.57	1685	0.43	3merge	8796	p < .000001.
Czech	Yes	12808	0.55	10391	0.45	3merge	54265	p < .000001.
Danish	No	954	0.27	2569	0.73	2merge	5011	p < .000001.
Dutch	No	3146.4	0.51	3042	0.49	3merge	6956	p < .000001.
English	No	1681	0.28	4264	0.72	2merge	12600	p < .000001.
Erzya	Yes	107	0.42	148	0.58	2merge	832	p < .000001.
Estonian	Yes	2440.9	0.27	6526	0.73	2merge	21290	p < .000001.
Finnish	Yes	952	0.29	2358	0.71	2merge	8661	p < .000001.
French	No	3864.5	0.51	3750	0.49	3merge	12708	p < .000001.
German	No	85216	0.66	44197	0.34	3merge	154653	p < .000001.
Greek	Yes	931	0.80	230	0.20	3merge	2382	p < .000001.
Hebrew	Yes	522	0.34	1036	0.66	2merge	3925	p < .000001.
Hungarian	Yes	517	0.61	334	0.39	3merge	1763	p = .000564.

[7] The binomial test gives us the probability of k successes (the number of the preferred merge system) in N independent trials (the number of objects, therefore transitive constructions), given a base probability p (the probability given by the null subject nature of the language) of an event. The binomial test (*z-test*) gives us the (one-tailed) probability of exactly the observed counts.

Indonesian	No	1451.1	0.50	1434	0.50	3merge	5795	$p < .000001$.
Irish	Yes	1043.8	0.62	648	0.38	3merge	4561	$p < .000001$.
Italian	Yes	3165.9	0.69	1444	0.31	3merge	10239	$p = .000004$
Italian Twitter	Yes	295	0.74	103	0.26	3merge	1222	$p < .000001$.
Korean	Yes	523	0.26	1501	0.74	2merge	23605	$p < .000001$.
Latvian	No	1186	0.29	2858	0.71	2merge	9780	$p < .000001$.
Lithuanian	Yes	153	0.25	464	0.75	2merge	2505	$p < .000001$.
Mbya Guarani	Yes	2	0.14	12	0.86	2merge	63	$p = .013168$.
Norwegian Bokmål	No	2591.8	0.28	6712	0.72	2merge	13904	$p = .000024$.
Norwegian Nynorsk	No	2342.3	0.24	7453	0.76	2merge	13218	$p < .000001$.
Persian	Yes	435	0.24	1393	0.76	2merge	3870	$p = .00004$.
Polish	Yes	1539	0.37	2596	0.63	2merge	15273	$p < .000001$.
Portoguese	Yes	3938.2	0.67	1914	0.33	3merge	11074	$p < .000001$.
Russian	No	4990.4	0.37	8532	0.63	2merge	33928	$p < .000001$.
Scottish Gaelic	No	131	0.18	608	0.82	2merge	1872	$p < .000001$.
Serbian	Yes	1045	0.59	738	0.41	3merge	3441	$p = .000549$.
Slovenian	Yes	1253	0.53	1114	0.47	3merge	7306	$p < .000001$.
Spanish	Yes	11633	0.73	4354	0.27	3merge	32061	$p < .000001$.
Swedish	No	706	0.22	2479	0.78	2merge	4241	$p < .000001$.
Thai	Yes	139	0.33	285	0.67	2merge	1734	$p < .000001$.
Turkish	Yes	649	0.39	1000	0.61	2merge	7402	$p < .000001$.
Upper Sorbian	Yes	94	0.57	71	0.43	3merge	368	$p = .00141$.
Uyghur	No	160	0.20	646	0.80	2merge	2301	$p < .000001$.
Vietnamese	Yes	514	0.33	1053	0.67	2merge	4078	$p < .000001$.
Warlpiri	Yes	4	0.09	41	0.91	2merge	50	$p < .000001$.
Wolof	Yes	319	0.18	1414	0.82	2merge	3319	$p < .000001$.
Yoruba	Yes	67	0.17	332	0.83	2merge	536	$p < .000001$.
Total		**190181**	**0.55**	**155351**	**0.45**	**3merge**	**692221**	$\boldsymbol{p} < .000001$*.

Table 2 – Language, whether the language is a null subject language establishing *p* (NULL P), frequency and distribution of 3-merge systems (FREQ3M, %3M), frequency and distribution of 2-merge systems (FREQ2M, %2M), preferred system representing the number of observation (K) for the binomial test, number of objects (OBJ) in the treebank representing the number of events (N) and the Z-TEST. * indicates that *p* has the same value for both 2- and 3-merge systems

There are no typological trends (e.g., German and Dutch prefer 3-merge systems, while Norwegian Bokmål and Norwegian Nynorsk show preference for 2-merge systems) and there is no correlation with the size of the corpus (r = 0.20). Considering all the languages as performance of a unique language, we can observe a marginal preference of 3-merge systems. This result is not indicative, since the distribution varies among languages having specific parametric values (null-subject vs. non-null-subject languages). Similarly, the two treebanks of Italian show relatively marginal different distributions (ISDT 0.69, TWITTIRO 0.74), but the linear regression model (r = 0.29, Poetry + 0.4303, Social Media + 0.3888, News +0.1957, Reviews +0.1667, Academic -0.2024, Nonfiction -0.2133, Medical -0.2899, Emails -0.6252) shows that there might be a trend. Indeed, registers like emails and (partially) academic/medical show a reduction of the usage of 3-merge systems. We leave the in-depth investigation of this result to further studies.[8]

The observed results in adult grammar open the question whether developmental paths in acquisition might show interesting dimensions of variation. Therefore, in section 5, we investigate child grammar (typical development and atypical development) to examine whether the complexity of the developing system gives rise to asymmetries.

5. Child Grammar and N-merge: some preliminary results

This section focuses on the frequency of 2-merge and 3-merge systems in typical development from observations in selected corpora of child grammar from Childes (MacWhinney 2000). In this study, we do not make any specific assumption on the structural configuration of child grammar with respect to early production of multiword utterances (see Guasti 2002: chapter 4 and reference therein). For the scope of our paper, we limit our analysis to the mere discussion of quantitative results of the extracted linguistic data adopting the model developed in section 4.

We restricted the investigation to two languages from two unrelated language families, namely English and Chinese. Following Guasti (2002: 101-310), we narrow our search on the period of time between two and four years (and around these two extremes).

5.1 Materials & Methods

After a manual analysis of the corpora collecting data from the languages under investigation, we decided to analyse longitudinal corpora from English and Mandarin Chinese. We investigated data extracted automatically from the *childesdb* (Sanchez *et al.* 2019). We performed the task on R (R development team 2016), isolating only target children's utterances in the relevant age (in terms of months). We only selected those sentences that are annotated for PoS.[9] A manual analysis was also conducted to evaluate the quality of the retrieved data. Some information on the size of the corpora, age range, the number of annotated utterances and references of the corpora under investigation are given in Table 3.

[8] These results are in line with the findings in Samo, Zhao and Gamhewage (2020) and Zhao *et al.* (2021), who have shown that syntactic complexity is cross-linguistically minimized in certain contexts, such as learning contents in public health with respect to social media, encyclopaedic entries and news.

[9] The morpho-syntactic annotation for corpora in Childes tendentially follows the guidelines provided by the annotation schemata MOR (doi:10.21415/T5B97X).

Language	Corpus	Age Range	Children	Annotated Utterances	References
English	Wells	1;6 – 5;0	32	17,964	Wells (1981)
Mandarin Chinese	Tong[A], Zhou3[B]	1;7-3;4[A] 0 :8 – 4 :5[B]	2	14,860	Deng and Yip (2015, 2018)[A]; Zhang and Zhou (2009)[B]

Table 3 – Relevant info on the corpora, age range, number of children, annotated utterances and related reference for the corpora under investigation

We translate the queries developed in section 4 according to the relevant annotation schemata. Both queries are based on the occurrences of patterns of labels in transitive construction given by a verbal element followed by an object. The morpho-syntactically annotated elements considered as 2-merge are personal pronouns (*pro:per, pro:sub*) and bare nominals (*n*); the combination of nominal elements with adjective (*adj*), articles (*det:art),* numerals *(num*) and classifiers (*cl*) were considered 3-merge. As discussed in section 4.1., we considered proper nouns (*n:prop*) as 3-merge elements. A manual analysis has been carried out to evaluate our semi-automatic retrieval. We summarize the queries in (4).

(4) 2-merge {pro:per}/{pro:sub}/{n} + transitive construction

3-merge {{det:art}/ {adj}/ {num} + {cl} + n} / {n:prop} + transitive construction

Naturally occurring sentences like (5a, b) will be labelled as 2-merge, while sentences like (5c) and (5d) as a 3 merge, some examples can be found in (5).

(5) a. English, 2-merge
I want my money (Elspeth, 2;6, ID: 9789443)

b. English, 3-merge
The dog have that ball (Abigail, 3:3, ID: 9761820)

c. Mandarin Chinese, 2-merge
我想画 个衣服 (Xue'er 1;11, ID: 5265930)

wo3	xiang3	hua4	ge4	yi1fu2
I	want	draw	cl	dress

d. Mandarin Chinese, 3-merge
这两个小朋友是男生 (Xue'er, 4;4 , ID: 5259789)

zhe4liang3ge4	xiao3peng2you3	shi4	nan2sheng1
this.two,cl	little-friends	are	boys

As for age of the target child (in months), we grouped the utterances in class intervals of one year (younger than 24 months, from 24 to 36 months, from 36 to 48 months, older than 48 months). Sub-section 5.2 summarizes the results.

5.2 Results

Results confirm H_2 (*the distributions of emergence of 2- and 3-merge systems in child grammar should differ*). Table 4 summarizes the results.

Age	EnUtt	En2mF	En%2m	En3mF	En%3m	ZhUtt	Zh2mF	Zh%2m	Zh3mF	Zh%3m
< 24	1244	20	0.016	2	0.002	2086	39	0.019	8	0.004
24 – 36	7119	551	0.077	52	0.007	6618	338	0.051	38	0.006
36 – 48	7069	783	0.111	79	0.011	4201	260	0.062	27	0.006
> 48	2532	352	0.139	30	0.012	1955	180	0.092	19	0.010

Table 4 – Age (in month), number of utterances in English and in Chinese (EnUtt, ZhUtt), frequency (F) of 2-merge and 3-merge in English and in Chinese (En2mF, En3mF, Zh2mF, Zh3mF) and distributions (%) of 2-merge and 3-merge in English and Chinese (En%2m, En%3m. Zh%2m, Zh%3m)

Our dataset shows that 2-merge system is the preferred option, in terms of distributions, in every age interval and in both languages. Similar cross-linguistic distributions can be observed before 24 months for both 2-merge (0.016 in English, 0.019 in Chinese) and 3-merge (0.002 in English and 0.004 in Chinese). After 48 months, we detect a comparably similar distribution for 3-merge configuration between English (0.012) and Chinese (0.010). The increase of the usage of these structures correlates with the age intervals: we can detect a correlation between age and the distribution of 2-merge ($r = 0.98$, $p < .05$) and 3-merge ($r = 0.97$, $p < .05$) in English. Asymmetries between the two systems are found Chinese: a slightly stronger correlation between age intervals is observed in 2-merge systems ($r = 0.99$, $p < .05$) compared to 3-merge systems ($r = 0.92$, $p = .08$).

Further research will involve a higher data set of languages and utterances (beyond longitudinal corpora) to detect more fine-grained differences.

5.3 Some notes on grammar in atypical development: a manual investigation

The last research question investigates grammars in children with atypical development. We extracted our data from the "Conti 2" corpus (Conti-Ramsden and Dykins 1991; Conti-Ramsden, Hutcheson and Grove 1995; Conti-Ramsden and Jones 1997), which contains transcripts, among others, from three children with specific language impairment (SLI). We first isolated the utterances of the relevant children (Andrew, Colin and Mark), then a manual investigation was carried out, due to the limited size of the dataset. This allowed us to detect every transitive construction and the nature of subjects, including null subjects (NS). Age intervals are different than the previous study in section 5.1. and 5.2., due to the nature of our dataset. We investigate utterances produced by the target children before 4 years, during the interval between 4 and 5 years and between 5 and 6 years, and after 6 years of age. Results are summarized in Table 5.

Age	Utterances	Verbs	Transitive	NS	Frequency 2m	Frequency 3m
<48	66	4	3	3	-	-
48-60	166	38	13	7	6	-

60-72	435	40	17	7	10	-
> 72	651	108	55	6	47	2

Table 5 – Age intervals in months, number of utterances, number of verbs, number of transitive constructions, number of null subjects (NS), frequency of 2-merge systems (2M) and frequency of 3-merge systems

In our (reduced) sample, we can observe that there is a strong preference for 2-merge systems in every age interval. As table 5 shows, 3-merge systems emerge later than 2-merge systems in our dataset. The two naturally occurring utterances are given in (6).

(6) a. a burglar he got it (Andrew, 78 months, ID: 14912223)
b. this one no beep the horn (Colin, 93 months, ID: 14916479)

From the transcriptions at our disposal, we can observe that (6a) might be a case of a topicalized subject or a configuration displaying a hanging topic. The example in (6b) involve a more complicated configuration with lack of agreement. Further research on bigger datasets is welcome to understand the dynamics concerning the relevant populations of speakers.

6. Conclusion

The results provided here shed light on the distribution of the typology of N-merge systems introduced in Rizzi (2016) adopting frequency as a dependent variable to test linguistic proposals (Merlo 2016 and related works).

We observed that in a rich dataset of 101 languages (102 treebanks), we can retrieve at least one occurrence of 3-merge in every variety. Such a result might confirm, at least for the sample we investigated, the prediction on the status of 3-merge systems in natural languages proposed in Rizzi (2016: 144).

Another important research question investigated focussed on preferences for 2-merge and 3-merge systems in adult grammar. We analysed 48 treebanks for 47 languages and we did not observe a clear typological trend (either null subject languages or language families). However, variability in terms of registers seems to be playing a role.

Developmental trends can be analysed via the exploration of child grammar corpora. We investigated two sets of longitudinal corpora in two unrelated languages (English and Chinese). Both languages display a clear pattern. Younger children from (around) 2 to (around) 4 years of age strongly prefer (in terms of frequency of production in spontaneous speech) 2-merge over 3-merge systems.

Along the same line, we investigated a corpus of utterances of children diagnosed with SLI in English. Though the reduced evidence at our disposal, we detected a trend: 2-merge systems seem to appear earlier with a higher frequency than 3-merge systems.

Future work should investigate in more fine-grained terms the correlation between frequency and complex structures with respect to specific register/genres. Finally, further research should also focus on all N-merge systems in acquisition by enlarging datasets and languages.

Acknowledgements

This research was supported by Science Foundation of Beijing Language and Culture University (supported by "the fundamental Research Funds for the Central Universities") 20YBB06. I am grateful to the audience of the *46th Incontro di Grammatica Generativa* (University of Siena) for useful remarks and questions. I would like to thank Yu Zhao and Emily Stanford for their precious comments on Chinese and English data.

References

Aepli, Noëmi, and Simon Clematide. 2018. "Parsing Approaches for Swiss German." In *Swisstext 2018: 3rd Swiss Text Analytics Conference* (Winterthur, Switzerland, 12-13 June 2018), vol. MMCCXXVI, *CEUR Workshop Proceedings*, ed. by Mark Cieliebak, Don Tuggener, and Fernando Benites, 6-16. <http://ceur-ws.org/Vol-2226/> (06/2021).

Agić, Željko, and Nikola Ljubešić. 2015. "Universal Dependencies for Croatian (that work for Serbian, too)." In *Proceedings of the 5th Workshop on Balto-Slavic Natural Language Processing* (Hissar, Bulgaria, 10-11 September 2015), ed. by Jakub Piskorski, Lidia Pivovarova, Jan Šnajder, Hristo Tanev, and Roman Yangarber, 1-8. <https://www.aclweb.org/anthology/W15-5301.pdf> (06/2021).

Ahrenberg, Lars. 2007. "LinES: An English-Swedish Parallel Treebank." In *Proceedings of The 16th Nordic Conference of Computational Linguistics (NODALIDA 2007)* (Tartu, Estonia, 25-26 May 2007), ed. by Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek, and Mare Koit, 270-273. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W07-2441> (06/2021).

Badmaeva, Elena, and Francis M. Tyers. 2017. "A Dependency Treebank for Buryat." In *Proceedings of the 15th International Workshop on Treebanks and Linguistic Theories (TLT15)* (Bloomington, Indiana, 20-21 January 2017), ed. by Markus Dickinson, Jan Hajič, Sandra Kübler, and Adam Przepiórkowski, 1-12. <http://ceur-ws.org/Vol-1779/01badmaeva.pdf> (06/2021).

Batchelor, Colin. 2019. "Universal Dependencies for Scottish Gaelic: Syntax." In *Proceedings of the Celtic Language Technology Workshop 2019* (Dublin, Ireland, 19 August 2019), ed. by Teresa Lynn, Delyth Prys, Colin Batchelor, and Francis Tyers, 7-15. European Association for Machine Translation. <https://www.aclweb.org/anthology/W19-6902.pdf> (06/2021).

Bejček, Eduard, Eva Hajičová, Jan Hajič, Pavlína Jínová, Václava Kettnerová, Veronika Kolářová, Marie Mikulová, Jiří Mírovský, Anna Nedoluzhko, Jarmila Panevová, Lucie Poláková, Magda Ševčíková, Jan Štěpánek, and Šárka Zikánová. 2013. *Prague dependency treebank 3.0.* LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics. Prague: Charles University. <http://hdl.handle.net/11858/00-097c-0000-0023-1aaf-3> (06/2021).

Belletti, Adriana. 2004. "Aspects of the Low IP Area." In *The Structure of CP and IP [The Cartography of Syntactic Structures, volume 2]*, ed. by Luigi Rizzi, 16-51.Oxford-New York, NY: Oxford UP.

Belletti, Adriana. 2015. "The Focus Map of Clefts: Extraposition and Predication." In *Beyond Functional Sequence [The Cartography of Syntactic Structures, volume 10]*, ed. by Ur Shlonsky, 42-60. Oxford-New York, NY: Oxford UP.

Belletti, Adriana, and Chris Collins. 2021. *Smuggling in Syntax*. Oxford: Oxford UP.

Belletti, Adriana, Naama Friedmann, Dominique Brunato, and Luigi Rizzi. 2012. "Does Gender Make a Difference? Comparing the Effect of Gender on Children's Comprehension of Relative Clauses in Hebrew and Italian." *Lingua* 122 (10): 1053-1069.

Belletti, Adriana, and Maria T. Guasti. 2015. *The Acquisition of Italian: Morphosyntax and its Interfaces in Different Modes of Acquisition*. Amsterdam-Philadelpia, PA: John Benjamins Publishing Company.

Belletti, Adriana, and Luigi Rizzi. 1981. "The Syntax of 'ne': Some Theoretical Implications." *The Linguistic Review* 1: 117-145.

Berthelot, Frédérique. 2017. *Movement of and out of Subjects in French*. Phd Dissertation, University of Geneva. <https://archive-ouverte.unige.ch/unige:96575> (06/2021).

Bhat, Irshad, Riyaz A. Bhat, Manish Shrivastava, and Dipti Sharma. 2018. "Universal Dependency Parsing for Hindi-English Code-Switching." In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* v*olume 1 (Long Papers)* (New Orleans, LA, 1-6 June 2018), ed. by Marilyn Walker, Heng Ji, and Amanda Stent, 987-998. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/N18-1090> (06/2021).

Bhat, Riyaz A., Rajesh Bhatt, Annahita Farudi, Prescott Klassen, Bhuvana Narasimhan, Martha Palmer, Owen Rambow, Dipti M. Sharma, Ashwini Vaidya, Sri Ramagurumurthy Vishnu, and Fei Xia. 2017. "The Hindi/Urdu Treebank Project." In *The Handbook of Linguistic Annotation*, ed. by Nancy Ide, and James Pustejovsky, 659-697.Dodrecht-New York, NY: Springer Press.

Bianchi, Valentina, and Mara Frascarelli. 2010. "Is Topic A Root Phenomenon?." *Iberia*: *An International Journal Of Theoretical Linguistics* 2 (1): 43-88.

Bianchi, Valentina, Giuliano Bocci, and Silvio Cruschina. 2015. "Focus Fronting and its Implicatures." In *Romance Languages and Linguistic Theory 2013: Selected papers from "Going Romance" Amsterdam 2013*, ed. by Enoch O. Aboh, Jeannette Schaeffer, and Petra Sleeman, 1-20. Amsterdam: John Benjamins.

Biberauer, Theresa, Ander Holmberg, Ian Roberts, and Michelle Sheehan. 2010. *Parametric Syntax: Null Subjects in Minimalist Theory.* Cambridge-New York, NY: Cambridge UP.

Bielinskienė, Agnė, Loïc Boizou, Jolanta Kovalevskaitė, and Erika Rimkutė. 2016. "Lithuanian Dependency Treebank Alksnis." In *Human Language Technologies – The Baltic Perspective*, ed. by Inguna Skadiņa, and Roberts Rozis, 107-114. Amsterdam: Ios Press. doi: 10.3233/978-1-61499-701-6-107.

Boeckx, Cedric, and Constantina Theofanopoulou. 2014. "A Multidimensional Interdisciplinary Framework for Linguistics: The Lexicon as a Key Study." *Journal of Cognitive Science* 15 (4): 403-420. doi: 10.17791/jcs.2014.15.4.403.

Bonan, Caterina. 2019. *On Clause-Internally Moved Wh-Phrases: Wh-To-Foc, Nominative Clitics, And The Theory Of Northern Italian Wh-In Situ*. PhD Dissertation, University of Geneva. <https://archive-ouverte.unige.ch/unige:119060> (06/2021).

Borer, Hagit. 1994. "The Projection of Arguments". *University of Massachusetts Occasional Papers in Linguistics* 20 (1): 17-49. <http://www.scholarworks.umass.edu/umop/vol20/iss1/3> (06/2021).

Borer, Hagit, and Kenneth Wexler. 1987. "The Maturation of Syntax." In *Parameter Setting. Studies in Theoretical Psycholinguistics*, ed. by Thomas Roeper, and Edwin Williams, 123-172. Dordrecht: Springer. doi: 10.1007/978-94-009-3727-7_6.

Borges Völker, Emanuel, Maximilian Wendt, Felix Hennig, and Arne Köhn. 2019. "HDT-UD: A Very Large Universal Dependencies Treebank for German." In *Proceedings Of The Third Workshop On Universal Dependencies (Udw, Syntaxfest 2019)* (Paris, France, 26-30 August 2019), ed. by Alexandre Rademaker, and Francis Tyers, 46-57. Cambridge, MA: MIT Press. <https://www.aclweb.Org/anthology/W19-8006> (06/2021).

Bosco, Cristina, Simonetta Montemagni, and Maria Simi. 2013. "Converting Italian Treebanks: Towards an Italian Stanford Dependency Treebank." In *Proceedings Of The 7th Linguistic Annotation Workshop And Interoperability With Discourse* (Sofia, Bulgaria, 8-9 August 2013), ed. by Antonio Pareja-Lora, Maria Liakata, and Stefanie Dipper, 61-69. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W13-2308/> (06/2021).

Bresnan, Joan, Shipra Dingare, and Christopher Manning. 2001. "Soft Constraints Mirror Hard Constraints: Voice And Person In English And Lummi." In *Proceedings of the LFG-01 Conference,* ed. by Miriam Butt, and Tracy H. King, 13-32. Stanford, CA: CSLI Publications.

Bross, Fabian. 2020. "Encoding Different Types Of Topics And Foci In German Sign Language. A Cartographic Approach To Sign Language Syntax." In *Glossa: A Journal of General Linguistics* 5 (1): 108. doi: 10.5334/Gjgl.1094.

Caloi, Irene. 2013. "The Comprehension Of Relative Clauses In Patients With Alzheimer's Disease." *Studies in Linguistics* 5: 5-24.

Cardinaletti, Anna. 2004. "Toward A Cartography Of Subject Positions." In *The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2*, ed. by Luigi Rizzi, 115-165. Oxford, New York, NY: Oxford UP.

Cecchini, Flavio M., Marco Passarotti, Paola Marongiu, and Daniel Zeman. 2018. "Challenges in Converting the Index Thomisticus Treebank into Universal Dependencies." In *Proceedings Of The Universal Dependencies Workshop 2018 (Udw 2018)* (Brussels, Belgium, 1 November 2018), ed. by Marie-Catherine de Marneffe, Teresa Lynn, and Sebastian Schuster, 27-36. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W18-6004.pdf> (06/2021).]

Çetinoğlu, Özlem, and Çagrı Çöltekin. 2019. "Challenges of Annotating a Code-Switching Treebank." In *Proceedings Of The 18th International Workshop On Treebanks And Linguistic Theories (TLT, Syntaxfest 2019)* (Paris, France, 28-29 August 2019), ed. by Marie Candito, Kilian Evang, Stephan Oepen, and Djamé Seddah, 82-90. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W19-7809.pdf> (06/2021).

Chesi, Cristiano. 2012. *Competence And Computation: Toward A Processing Friendly Minimalist Grammar*. Padova: Unipress.

Chesi, Cristiano. 2015. "On Directionality Of Phrase Structure Building." *Journal of Psycholinguistic Research* 44 (1): 65-89.

Chesi, Cristiano, and Paolo Canal. 2019. "Person Features and Lexical Restrictions in Italian Clefts." In *Frontiers in Psychology* 10. doi: 10.3389/fpsyg.2019.02105.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1981. *Lectures On Government And Binding: The Pisa Lectures*. Dordrecht: Foris Publications.

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

Chomsky, Noam. 2013. "Problems Of Projection." *Lingua* 130: 33-49.

Cignarella, Alessandra T., Cristina Bosco, Viviana Patti, and Mirko Lai. 2018. "Application and Analysis of a Multi-layered Scheme for Irony on the Italian Twitter Corpus TWITTIRÒ." In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki, Japan, 7-12 May 2018), ed. by Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga, 4204-4211. Paris : European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L18-1664.pdf> (06/2021).

Cinque, Guglielmo. 2014. *Typological Studies: Word Order and Relative Clauses*. New York, NY: Routledge.

Collins, Chris. 2005. "A Smuggling Approach To Raising In English." *Linguistic Inquiry* 36 (2): 289-298.

Conti-Ramsden, Gina, and Jane J. Dykins. 1991. "Mother-Child Interactions with Language-Impaired Children and their Siblings." *British Journal Of Disorders Of Communication* 26 (3): 337-354.

Conti-Ramsden, Gina, Graeme D. Hutcheson, and John Grove. 1995. "Contingency And Breakdown: Specific Language Impaired Children's Conversations With Their Mothers And Fathers." *Journal Of Speech And Hearing Research* 38 (6): 1290-1302. doi: 10.1044/jshr.3806.1290.

Conti-Ramsden, Gina, and Melanie Jones, 1997. "Verb Use in Specific Language Impairment." *Journal Of Speech And Hearing Research* 40 (6): 1298-1313. doi: 10.1044/jslhr.4006.1298.

Dehaene, Stanislas, Florent Meyniel, Catherine Wacongne, Liping Wang, and Christophe Pallier. 2015. "The Neural Representation Of Sequences: From Transition Probabilities To Algebraic Patterns And Linguistic Trees." *Neuron* 88 (1): 2-19.

Deng, Xiangjun, Ziyin Mai, and Virginia Yip. 2015. *An Aspectual Account Of Ba And Bei Constructions In Child Mandarin.* Paper Presented At The International Symposium On Psycholinguistics Of Second Language Acquisition And Bilingualism, Chinese University Of Hong Kong.

Deng, Xiangjun, and Virginia Yip, 2018. "A Multimedia Corpus Of Child Mandarin: The Tong Corpus." *Journal Of Chinese Linguistics* 46 (1): 69-92.

De Freitas, Marília F.P. 2017. *A Posse Em Apurinã: Descrição De Construções Atributivas E Predicativas Em Comparação Com Outras Línguas Aruák*. PhD Dissertation. Parà: Federal University of Pará.

Di Domenico, Elisa, Ioli Baroncini, and Andrea Capotorti. 2020. "Null and Overt Subject Pronouns in Topic Continuity and Topic Shift: An Investigation of the Narrative Productions of Italian Natives, Greek Natives and Near-native Second Language Speakers of Italian with Greek as a First language." *Glossa: A Journal Of General Linguistics* 5 (1): 117. doi: 10.5334/gjgl.1009.

Droganova, Kira, Olga Lyashevskaya, and Daniel Zeman. 2018. "Data Conversion And Consistency Of Monolingual Corpora: Russian Ud Treebanks." In *Proceedings Of The 17th International Workshop On Treebanks And Linguistic Theories (Tlt 2018)*, (Oslo, Norway, 13-14 December 2018), ed. by Dag Haug, Stephan Oepen, Lilja Øvrelid, Marie Candito, and Jan Hajic, 52-65. Linköping: Linköping University Electronic Press.

Dryer, Matthew S. 2009. "Problems Testing Typological Correlations with the Online WALS." *Linguistic Typology* 13 (1): 121-135. doi: 10.1515/lity.2009.007.

Dryer, Matthew S. 2013. "Expression Of Pronominal Subjects." In *The World Atlas Of Language Structures Online*, ed. by Matthew S. Dryer, and Martin Haspelmath. <http://wals.info/chapter/101> (06/2021).

Dryer, Matthew S., and Martin Haspelmath. 2013. *WALS Online*. Leipzig: Max Planck Institute For Evolutionary Anthropology. <https://wals.info/> (06/2021).

Durrleman, Stephanie, Loyse Hippolyte, Sandrine Zufferey, Katia Iglesias, and Nouchine Hadjikhani. 2015. "Complex Syntax in Autism Spectrum Disorders: A Study of Relative Clauses." *International Journal Of Language and Communication Disorders* 50 (2): 260-267.

Eckhoff, Hanne M., and Aleksandrs Berdicevskis. 2015. "Linguistics vs. Digital Editions: The Tromsø Old Russian and OCS Treebank." *Scripta and E-Scripta* 14-15: 9-25.

Ephrem Seyoum, Binyam, Yusuke Miyao, and Baye Yimam. 2018. "Universal Dependencies For Amharic." In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki, Japan, 7-12 May 2018), ed. by Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga, 2216-2222. Paris: European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L18-1350> (06/2021).

Frascarelli, Mara. 2007. "Subjects, Topics and the Interpretation of Referential Pro." *Natural Language and Linguistic Theory* 25 (4): 691-734.

Frascarelli, Mara, and Marco Casentini. 2019. "The Interpretation of Null Subjects in a Radical Pro-drop Language: Topic Chains and Discourse-semantic Requirements in Chinese." *Studies In Chinese Linguistics* 40 (1): 1-45. doi: 10.2478/scl-2019-0001.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. "Types of Topics in German and Italian." In *On Information Structure, Meaning and Form: Generalizations Across Languages*, ed. by Kerstin Schwabe, and Susanne Winkler, 87-116. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company.

Frauenfelder, Ulrich, Juan Segui, and Jacques Mehler. 1980. "Monitoring Around the Relative Clause." *Journal of Verbal Learning and Verbal Behavior* 19 (3): 328-337.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. "Relativized Relatives: Types of Intervention in the Acquisition of A-bar Dependencies." *Lingua* 119 (1): 67-88.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2020. "Growing Trees. The Acquisition of the Left Periphery." Manuscript. University of Tel Aviv, University of Siena, Collège de France. <https://ling.auf.net/lingbuzz/005369> (06/2021).

Gallego, Ángel J. 2008. "Four Reasons to Push Down the External Argument." Manuscript. Universitat Autònoma De Barcelona. <http://filcat.uab.cat/clt/membres/professors/gallego/pdf/external_argument.pdf.> (06/2021).

Gökirmak, Memduh, and Francis M. Tyers. 2017. "A Dependency Treebank for Kurmanji Kurdish." In *Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)* (Pisa, Italy, 18-20 September 2017), ed. by Simonetta Montemagni, and Joakim Nivre, 64-72. Linköping: Linköping University Electronic Press. <https://www.aclweb.org/anthology/W17-6509.pdf> (06/2021).

Grillo, Nino. 2008. *Generalized Minimality: Sinctactic Underspecification in Broca's Aphasia*. PhD Dissertation. York: University of York.

Guasti, Maria T. 2002. *Language Acquisition – The Growth of Grammar*. Cambridge, MA: MIT Press.

Guasti, Maria T. 2017. *Language Acquisition – The Growth of Grammar*. Cambridge, MA: MIT Press.

Gulordava, Kristina, and Paola Merlo. 2020. "Computational Quantitative Syntax: The Case Of Universal 18." In *Romance Languages And Linguistic Theory 16: Selected Papers From The 47th Linguistic Symposium On Romance Languages (LSRL)*, *Newark, Delaware,* ed. by Irene Vogel, 110-132. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company. doi: 10.1075/rllt.16.08gul.

Haegeman, Liliane. 1990. "Non-Overt Subjects in Diary Contexts." In *Grammar in Progress: Glow Essays for Henk Van Riemsdijk,* ed. by Joan Mascaró, and Marina Nespor, 167-174. Dodricht: Foris Publications.

Hale, Ken, and Samuel J. Keyser. 1998. "The Basic Elements of Argument Structure." In *MIT Working papers in linguistics,* vol. XXXII, *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, ed. by Heidi Harley. 73-118. Cambridge, MA: MIT Press.

Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. "The Weka Data Mining Software: An Update." *ACM SGKDD Explorations Newsletter* 11 (1): 10-18. doi: 10.1145/1656274.1656278.

Haug, Dag T., and Marius L. Jøhndal. 2008. "Creating a Parallel Treebank of the Old Indo-European Bible Translations." In *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)* (Marrakech, Morocco, 1 June 2008), ed. by Caroline Sporleder and Kiril Ribarov, 27-34. Cambridge, MA: MIT Press.

Hellwig, Oliver, Salvatore Scarlata, Elia Ackermann, and Paul Widmer. 2020. "The Treebank of Vedic Sanskrit." In *Proceedings of the 12th Language Resources and Evaluation Conference* (Marseille, France, 11-16 May 2020), ed. by Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, 5137-5146. Paris: European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/2020.lrec-1.632.pdf> (06/2021).

Heinecke, Johannes, and Francis M. Tyers. 2019. "Development of a Universal Dependencies treebank for Welsh." In *Proceedings of the Celtic Language Technology Workshop* (Dublin, Ireland, 19 August 2019), ed. by Teresa Lynn, Delyth Prys, Colin Batchelor, and Francis Tyers, 21-31. European Association for Machine Translation. <https://www.aclweb.org/anthology/W19-6904.pdf> (06/2021).

Holmberg, Anders, and Ian Roberts. 2013. "The Syntax-Morphology Relation." *Lingua* 130: 111-131. doi: 10.1016/j.lingua.2012.10.006.

Ibbotson, Paul. 2013. "The Scope of Usage-Based Theory." *Frontiers in Psychology* 4. doi:10.3389/fpsyg.2013.00255.

Jaeggli, Osvaldo, and Kenneth J. Safir. 1989. "The Null Subject Parameter and Parametric Theory." In *The Null Subject Parameter*, ed. by Osvaldo Jaeggli, and Kenneth J. Safir, 1-44. Dordrecht-Boston, MA-London: Kluwer Academic Publisher.

Johannsen, Anders, Héctor Martínez Alonso, and Barbara Plank. 2015. "Universal Dependencies For Danish." In *Proceedings of The Fourteenth Internationall Workhop on Treebank and Linguistic Theories (TLT14)* (Warsaw, Poland, 11-12 December 2015), ed. by Markus Dickinson, Erhard Hinrichs, Agnieszka Patejuk, and Adam Przepiórkowski, 157-167. <http://tlt14.ipipan.waw.pl/proceedings/> (06/2021).

Kam, Carla H., and Elissa L. Newport. 2005. "Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change." *Language Learning and Development* 1 (2): 151-195.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda, and Dominique Sportiche. 1991. "The Position of Subjects" *Lingua* 85 (2-3): 211-258. doi: 10.1016/0024-3841(91)90022-W. <https://www.sciencedirect.com/science/article/pii/002438419190022W> (06/2021).

Larson, Richard K. 1988. "On the Double Object Construction." *Linguistic Inquiry* 19 (3): 335-391.

Longobardi, Giuseppe. 1994. "Reference and Proper Names: a Theory of N-Movement in Syntax and Logical Form." *Linguistic Inquiry* 25 (4): 609-665.

Luukko, Mikko, Aleksi Sahala, Sam Hardwick, and Krister Lindén. 2020. "Akkadian Treebank for Early Neo-Assyrian Royal Inscriptions." In *Proceedings of the 19th International Workshop on Treebanks*

and Linguistic Theories – Düsseldorf (Düsseldorf, Germany, 27-28 October, 2020), ed. by Kilian Evang, Laura Kallmeyer, Rafael Ehren, Simon Petitjean, Esther Seyffarth, and Djamé Seddah, 124-134. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/2020.tlt-1.11> (06/2021).

MacWhinney, Brian. 2000a. *The CHILDES project: Tools for Analyzing Talk: Volume I: Transcription Format and Programs*. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, Brian. 2000b. *The CHILDES Project: Tools for Analyzing Talk: Volume II: The Database*. Mahwa, NJ: Lawrence Erlbaum Associates.

Makazhanov, Aibek, Aitolkyn Sultangazina, Olzhas Makhambetov, and Zhandos Yessenbayev. 2015. "Syntactic Annotation of Kazakh: Following the Universal Dependencies Guidelines. A Report." In *Proceedings of the International Conference "Turkic Language Processing". TurkLang-2015* (Kazan, Russia, 17-19 September 2015), 338-350. Kazan: Tartan Academy of Sciences. <http://www.turklang.net/wp-content/uploads/2017/05/proceedings.pdf > (06/2021).

Martini, Karen, Adriana Belletti, Santi Centorrino, and Maria Garraffa. 2020. "Syntactic Complexity in the Presence of an Intervener: The Case of an Italian Speaker with Anomia." *Aphasiology* 34 (8): 1016-1042. doi: 10.1080/02687038.2019.1686744.

Maslinsky, Kirill. 2014. "Daba: a Model and Tools for Manding Corpora." In *Proceedings Of Talaf 2014: Traitement Automatique Des Langues Africaines* (Marseille, France, 1 July 2014), ed. by Mathieu Mangeot, and Fatiha Sadat, 114-122. Association pour le Traitement Automatique des Langues. <https://www.aclweb.org/anthology/W14-6502> (06/2021).

Merlo, Paola. 2016. "Quantitative Computational Syntax: Some Initial Results." *Italian Journal Of Computational Linguistics* 2 (1): 11-30.

Merlo, Paola, and Sarah Ouwayda. 2018. "Movement and Structure Effects on Universal 20 Word Order Frequencies: A Quantitative Study." *Glossa: A Journal of General Linguistics* 3 (1): 84. doi:10.5334/gjgl.149.

Merlo, Paola, and Suzanne S. Stevenson. 1998. "What Grammars Tell Us About Corpora: The Case of Reduced Relative Clauses". In *Proceedings of the Sixth Workshop on Very Large Corpora* (Montreal, Canada, 15-16 August 1998), ed. by Eugene Charniak, 134-142. Montreal: University of Montreal. <https://www.aclweb.org/anthology/W98-1116> (06/2021).

Moro, Andrea. 1997. "Dynamic Antisymmetry: Movement as a Symmetry-breaking Phenomenon." *Studia Linguistica* 51 (1): 50-76.

Murphy, Elliot. 2015. "The Brain Dynamics of Linguistic Computation." *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.01515.

Neeleman, Ad, and Kriszta Szendrői. 2007. "Radical Pro Drop and the Morphology of Pronouns." *Linguistic Inquiry* 38 (4): 671-714.

Nivre, Joakim. 2015. "Towards a Universal Grammar for Natural Language Processing." In *International Conference on Intelligent Text Processing and Computational Linguistics: 16th International Conference, CICLing 2015, Proceedings, Part I ,* (Cairo, Egypt, April 14-20, 2015), ed. by Alexander Gelbukh, 3-16. Cham: Springer.

Ojha, Atul Kr., and Daniel Zeman. 2020. "Universal Dependency Treebanks for Low-Resource Indian Languages: The Case of Bhojpuri." In *Proceedings of the WILDRE5 – 5th Workshop on Indian Language Data: Resources and Evaluation* (Marseille, France, 11-16 May 2020), ed. by Girish Nath Jha, Kalika Bali, Sobha L., S.S. Agrawal, and Atul Kr. Ojha, 33-38. Paris: European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/2020.wildre-1.7> (06/2021).

Partanen, Niko, Rogier Blokland, KyungTae Lim, Thierry Poibeau, and Michael Rießler. 2018. "The First Komi-Zyrian Universal Dependencies Treebanks." In *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)* (Brusselles, Belgium, 1 November 2018), ed. by Marie-Catherine de Marneffe, Teresa Lynn, and Sebastian Schuster, 126-132. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W18-6015> (06/2021).

Perlmutter, David M. 1978. "Impersonal Passives and the Unaccusative Hypothesis." In *Proceedings of the 4th Annual Meeting of The Berkeley Linguistics Society* (BLS) 4: 157-189. University of California: Escolarship. <https://escholarship.org/uc/item/73h0s91v> (06/2021).

Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP". *Linguistic Inquiry* 20 (3): 365-424.

Preminger, Omer. 2008. "Argument Externality." Manuscript. <https://ling.auf.net/lingbuzz/000569> (06/2021).

Prokopidis, Prokopis, and Haris Papageorgiou. 2017. "Universal Dependencies for Greek." In *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)* (Gohenburg, Sweden, 22 May 2017), ed. by Marie-Catherine de Marneffe, Joakim Nivre, and Sebastian Schuster, 102-106. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W17-0413> (06/2021).

R Development Core Team. 2016. *R: A Language And Environment For Statistical Computing*. Vienna: R Foundation For Statistical Computing.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge UP.

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

Rizzi, Luigi. 1997. "The Fine Structure Of The Left Periphery." In *Elements Of Grammar: Handbook of Generative Syntax*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2006. "On the Form of Chains: Criterial Positions and ECP Effects." In *Current Studies in Linguistics*, vol. XLII, *Wh-Movement: Moving On*, ed. by Lisa Cheng, and Norbert Corver, 97-133. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2015a. "Cartography, Criteria, And Labeling." In *Beyond Functional Sequence [The Cartography Of Syntactic Structures, Volume 10]*, ed. by Ur Shlonsky, 314-338. Oxford-New York, NY: Oxford UP. doi: 10.1093/acprof:oso/9780190210588.003.0017.

Rizzi, Luigi 2015b. "Notes on Labeling and Subject Positions." In *Structures, Strategies and Beyond – Studies In Honour Of Adriana Belletti*, ed. by Cornelia Hamann, Elisa Di Domenico, and Simona Matteini, 17-46. Amsterdam: John Benjamins Publishing Company.

Rizzi, Luigi. 2016. "Monkey Morpho-Syntax and Merge-Based Systems." *Theoretical Linguistics* 42 (1-2): 139-145. doi: 10.1515/tl-2016-0006.

Rizzi, Luigi, and Giuliano Bocci. 2017. "Left Periphery of the Clause: Primarily Illustrated for Italian." In *The Blackwell Companion To Syntax*, ed. by Martin Everaert, and Henk Van Riemsdijk, 1-30. Hoboken, NJ: John Wiley and Sons.

Rueter, Jack. 2018. *Erme UD Moksha (Version v1.0)*. doi:10.5281/zenodo.1156112.

Rueter, Jack, Niko Partanen, and Larisa Ponomareva. 2020. "On the Questions in Developing Computational Infrastructure for Komi-Permyak." In *Proceedings of the Sixth International Workshop on Computational Linguistics of Uralic Languages*,(Wien, Austria, 10-11 January 2020), ed. by Tommi A Pirinen, Francis M. Tyers, and Michael Rießler, 15-25. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/2020.iwclul-1.3> (06/2021).

Rueter, Jack, and Francis M. Tyers. 2018. "Towards an Open-source Universal-dependency Treebank for Erzya." In *Proceedings of the Fourth International Workshop on Computational Linguistics of Uralic Languages* (Helsinki, Finland, 8-9 January 2018), ed. by Tommi A. Pirinen, Michael Rießler, Jack Rueter, Trond Trosterud, and Francis M. Tyers, 106-118. Cambridge, MA: MIT Press <https://www.aclweb.org/anthology/W18-0210> (06/2021).

Samo, Giuseppe. 2019. "Cartography and Locality in German: a Quantitative Study with Dependency Structures." *Rivista Di Grammatica Generativa/Research in Generative Grammar* 41 (5): 1-26. <http://lear.unive.it/jspui/handle/11707/7782> (06/2021).

Samo, Giuseppe, and Paola Merlo. 2019. "Intervention Effects in Object Relatives in English and Italian: A Study in Quantitative Computational Syntax." In *Proceedings of the First Workshop on Quantitative Syntax (Quasy, Syntaxfest 2019)* (Paris, France, 26-30 August 2019), ed. by Xinying Chen, and Ramon Ferrer-i-Cancho, 46-56. Cambridge, MA: MIT Press. doi: 10.18653/v1/w19-7906.

Samo, Giuseppe, Yu Zhao, and Gaya Gamhewage. 2020. "Syntactic Complexity of Learning Content in Italian for Covid-19 Frontline Responders: A Study on WHO's Emergency Learning Platform." *Verbum* 11. doi:10.15388/Verb.15.

Sanchez, Alessandro, Stephan C. Meylan, Mika Braginsky, Kyle E. MacDonald, Daniel Yurovsky, and Michael C. Frank. 2019. "Childes-db: A Flexible and Reproducible Interface to the Child Language Data Exchange System." *Behavior Research Methods* 51: 1928-1941.

Schlenker, Philippe, Emmanuel Chemla, Anne M. Schel, James Fuller, Jean-Pierre Gautier, Jeremy Kuhn, Dunja Veselinović, Kate Arnold, Cristiane Cäsar, Sumir Keenan, Alban Lemasson, Karim Ouattara, Robin Ryder, and Klaus Zuberbühler. 2016. "Formal Monkey Linguistics." *Theoretical Linguistics* 42 (1-2): 1-90. doi: 10.1515/tl-2016-0001.

Shim, Ji Y. 2016. "Mixed Verbs in Code-Switching: The Syntax of Light Verbs." *Languages* 1 (1): 8. doi: 10.3390/languages1010008.

Si, Fuzhen. 2019. "A Cartographic Study of Light Verb Constructions." *Yuwen Xuexi (Literacy Study)*1: 1-20.

Stanford, Emily N. 2020. *The Language-Cognition Interface: Executive Functions and Syntax in Atypical Development.* PhD Dissertation, University of Geneva. doi: 10.13097/archive-ouverte/unige:144700.

Prévost, Sophie, and Achim Stein. 2013. "Syntactic Annotation of Medieval Texts: the Syntactic Reference Corpus of Medieval French (SRCMF)." In *New Methods in Historical Corpora: Corpus Linguistics and International Perspectives on Language*, ed. by Paul Bennett, Martin Durrell, Silke Scheible, and Richard J. Whitt, 275-282. Tübingen: Gunter Narr Verlag.

Tsarfaty, Reut. 2013. "A Unified Morpho-Syntactic Scheme of Stanford Dependencies." In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (Sofia, Bulgaria, 4-9 August 2013), ed. by Hinrich Schuetze, Pascale Fung, and Massimo Poesio, 578-584. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/P13-2103> (06/2021).

Tyers, Francis M., and Vinit Ravishankar. 2018. "A Prototype Dependency Treebank for Breton." In *Actes de la Conférence TALN* (Rennes, France, 14-18 May 2018)*, vol 1 − Articles longs, articles courts de TALN*, ed. by Pascale Sébillot, and Vincent Claveau, 197-204. Rennes: ATALA. <https://www.aclweb.org/anthology/2018.jeptalnrecital-court.1> (06/2021).

Tyers, Francis M., and Mariya Sheyanova. 2017. "Annotation Schemes in North Sámi Dependency Parsing." In *Proceedings of the Third Workshop on Computational Linguistics for Uralic Languages* (St. Petersburg, Russia, 23-24 January 2017), ed. by Francis M. Tyers, Michael Rießler, Tommi A. Pirinen, and Trond Trosterud, 66-75. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W17-0607 > (06/2021).

Tyers, Francis M., Mariya Sheyanova, Aleksandra Martynova, Pavel Stepachev, and Konstantin Vinogorodskiy. 2018. "Multi-source Synthetic Treebank Creation for Improved Cross-lingual Dependency Parsing." In *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)* (Brussels, Belgium, 1 November 2018), ed. By Marie-Catherine de Marneffe, Teresa Lynn, and Sebastian Schuster, 144-150. Cambridge, MA: MIT Press. <https://www.aclweb.org/anthology/W18-6017> (06/2021).

Türk, Utku, Furkan Atmaca, Şaziye B. Özateş, Gözde Berk, Seyyit T. Bedir, Abdullatif Köksal, Balkız Öztürk Başaran, Tunga Güngör, and Arzucan Özgür. 2020. *Resources For Turkish Dependency Parsing: Introducing The Boun Treebank And The Boat Annotation Tool*. Manuscript. <https://arxiv.org/abs/2002.10416 (06/2021).

Velldal, Erik, Lilja Øvrelid, and Petter Hohle. 2017. "Joint UD Parsing of Norwegian Bokmål and Nynorsk." In *Proceedings of the 21st Nordic Conference on Computational Linguistics* (Gothenburg, Sweden, 22-24 May 2017), ed. by Jörg Tiedemann, 1-10. Linköping: Linköping University Electronic Press. <https://www.aclweb.org/anthology/W17-0201.pdf > (06/2021).

Volk, Martin, Johannes Graën, and Elena Callegaro. 2014. "Innovations in Parallel Corpus Search Tools." In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (Reykjavik, Iceland, 26-31 May 2014), ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, 1-7. Paris: European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2014/pdf/504_Paper.pdf.> (06/2021).

Wells, Gordon. 1981. *Learning Through Interaction: The Study of Language Development*. Cambridge: Cambridge UP. doi: 10.1017/CBO9780511620737.

Yang, Charles. 2013. "Ontogeny and Phylogeny of Language." *Proceedings of the National Academy of Sciences* 110 (16): 6324-6327.

Yang, Charles. 2015. *Generalization and Probability Matching*. Manuscript. University of Pennsylvania. <https://www.ling.upenn.edu/~ycharles/probmatch.pdf> (06/2021).

Yavrumyan, Marat M. 2019. "Universal Dependencies for Armenian." (Paris, France, 3-5 October 2019). In *International Conference on Digital Armenian*. Paris: Institut National des Langues et Civilisations Orientales (INALCO).

Zeman, Daniel. 2017. "Slovak Dependency Treebank in Universal Dependencies." *Jazykovedný Časopis* 68 (2): 385-395.

Zeman, Daniel, Joakim Nivre, and Mitchell M. Abrams, *et al.* 2020. *Universal Dependencies 2.7*. LINDAT/CLARIAH-CZ Digital Library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics. Prague: Charles University. <http://hdl.handle.net/11234/1-3424> (06/2021).

Zhang, Li, and Jing Zhou. 2009. "The Development of Mean Length of Utterance in Mandarin-speaking Children (汉语儿童平均语句长度发展研究)." In *The Application and Development of International Corpus-based Research Methods* (汉语儿童语言发展研究: 国际儿童语料库研究方法的应用与发展), ed. by Jing Zhou, 40-58. Beijing: Education Science Publishing House (教育科学出版社).

Zhao, Yu, Giuseppe Samo, Heini Utunen, Oliver Stucke, and Gaya Gamhewage. 2021. "Evaluating Complexity of Digital Learning in a Multilingual Context: A Cross-Linguistic Study on WHO's Emergency Learning Platform". In *Public Health and Informatics*, vol. CCLXXXI of *Studies of Health Technologies and Informatics* ed. by John Mantas, Lăcrămioara Stoicu-Tivadar, Catherine Chronaki, Arie Hasman, Patrick Weber, Parisis Gallos, Mihaela Crișan-Vida, Emmanouil Zoulias, and Oana S. Chirila, 516-517. Amsterdam: IOS Press. doi: 10.3233/SHTI210222.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

OPEN ACCESS

Citation: Tesfay Tewolde Yohannes (2021) Gender and Number Morphology in Ethio-Eritrean Semitic Languages. *Qulso* 7: pp. 131-154. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12006

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Gender and number morphology in Ethio-Eritrean semitic languages

Tesfay Tewolde Yohannes
University of Firenze (<tesfay.tewoldeyohannes@unifi.it>)
University of Napoli, College of Arts and Social Sciences (Eritrea), University of Addis Abeba

Abstract:

The vowel -ā can be identified as a marker of nominal and verbal plurality in different Semitic and Afroasiatic languages. The vowels -ā (feminine plural) and -ū (masculine plural) which are used for both internal and external plurals are, according to Hasselbach (2007), derived from a verbal system and a predicative adjective. In Semitic languages, ā, ū and *at* (< *a+t*) mark gender and/or number. They are interdependent and can be used as classifiers. Based on the analysis of data from the languages in question, this article argues the Semitic gender and/or number markers indicated above and their reflexes *-ačč*, *-očč*, *-o* and *-ot* can be used as gender and/or number markers in Ethio-Eritrean Semitic languages. As in other Semitic languages, gender and number markers can function as classifiers in EES languages too. Even though there is interdependence of gender with number, the use of the former on verbal and nominal forms may be regarded as more original than the latter. However, more research is also needed to explore the diachronic relationship between them.

Keywords: *Classifier, Gender, Number, Template*

1. *Introduction*

Rubin (2005) argues the participle is properly a nominal form, historically inflected for number and gender. Furthermore, Rubin says verbal noun or verbal adjective developed into a stative verb in Proto-Afro-Asiatic. In several world languages, the development of a passive form into a perfective verb can be observed (cf. Kouwenberg 2006 among others). In Ancient Hebrew, verbs can be used to express concepts which English expresses with adjectives (cf. Steiner 1997: 155). In Aramaic, Kaufman (1997: 124) says "adjectives probably were originally limited to the passive participles and the related form *kattīb*". In Neo-Aramaic, the verbal base is derived from an active or passive old participle (cf. Jastrow 1997: 360). As indicated in Lipiński

(1997), Semitic languages have nominal sentences consisting of two nominal phrases (with no copula) that may correspond roughly to English sentences containing "is". The predicate of a nominal sentence syntactically assumes a quasi-verbal function. In such nominal sentences, the predicate generally follows the subject as in *Yohannis mäkonnin* 'Yohannes (is) a/the judge' in Gɨʕɨz, *Adad* šarrum 'Adad (is) king' in Old Babylonian, *Yhwh roʕi* 'Yawhe (is) my shepherd' in Hebrew. If there is the need to express emphasis on the predicate, however, the word order can be inverted and the predicate may come in front of the sentence as in the case of Hebrew ʿāpār ʔattā 'dust you (are)' (cf. Lipiński 1997: 484-485). Carver (2016) says (a) the nominal origin of Akkadian stative is unquestionable (b) the morphological base of the Akkadian stative is undeniably related to verbal adjective base (c) the Akkadian stative is a non-finite verb morphologically marked for gender, number and person.

According to Hasselbach (2007: 132-135), the Semitic external plural and dual markers -ā, -āt and -ū can be derived from the verbal system and the predicative adjective. Hasselbach argues the nominal feminine plural -āt can be derived from the predicative feminine plural -ā by the suffixation of the feminine singular marker *t*.

In Semitic languages, we can observe the relationship among person, gender and number markers in demonstratives, independent pronouns, possessive suffixes attached to nouns and pronominal affixes attached to verbs etc. It is indicated in the literature that Semitic languages have verbal affixes, possessive suffixes, and affixes attached to demonstratives and independent pronouns[1] with a number marker, primary gender markers and secondary markers. Semitic languages have *-a* for the masculine and *-i* for the feminine as primary gender markers. Moreover, the Semitic languages have also *-ā* (as a secondary feminine gender marker) and -ū (as a secondary masculine gender marker), while *n* (which may become *m* in the masculine) functions as a number marker. According to Buccellati (1996) and others, the primary and secondary gender markers are represented by short and long vowels respectively.

In Semitic languages as in Akkadian, we also see that the originally secondary gender markers *-ū* and -ā can be used as number markers of verbs (cf. Buccellati 1996, 1997). In fact, the originally secondary gender markers *-ā* and -ū in verbs are related to external and internal plurals of nouns. Semitic languages use the gender markers *a*, *u, t* and a number marker *n* to indicate plurality. In Hasselbach (2007: 123, 129), we see *-au* in Ancient Egyptian, *-aw* in Middle Egyptian and *-aw* in Berber to mark masculine plural.

The article discusses Ethio-Eritrean Semitic (EES) gender and number markers. The primary goal of the work is to explore the relationship between gender and number markers in the languages in question. It focuses on data from Tigrinya, Amharic and Tigre. However, it also examines data from Gɨʕɨz, Harari and other EES languages. The article is organized as follows. Section 1 deals with the introduction. Section 2 offers an overview of EES number and gender markers. Section 3 concerns the relationship among EES gender and number markers. Section 4 discusses the role of EES gender and/or number markers as classifiers. Section 5 deals with issues regarding the "multi-plurals". Section 6 concerns the position of agreement morphemes on the structure. Section 7 provides a conclusion.

[1] In the literature, we can observe that demonstratives can develop into definite articles (cf. Lyons 1999 among others), while pronouns and definite articles may occupy the same position in the structure. In fact, we can observe in languages such as Ugaritic, Chaha and Sabaic that the same word (or similar words) may indicate a demonstrative and a pronoun (cf. Lipiński 1997; Pardee 1997; Tesfay 2016 among others). As indicated in Fuß (2005: 2-5), verbal agreement markers can be derived from (originally independent) personal pronouns (cf. also Simpson 2009).

2. An Overview on EES Number and Gender Morphemes

In EES (Ethio-Eritrean Semitic) and other Semitic languages, gender distinctions can be observed as it affects the forms of the related words. We can see in the literature that number and gender markers are related. This section offers an overview of number and gender markers in EES languages.

Nouns in EES and other Semitic languages can have singular and plural forms and the latter can be divided into internal and external plurals. In the literature, it is indicated that internal plurals of nouns are related to internal plurals of verbs (cf. Greenberg 1955, 1991; Benmamoun 2003; Tesfay 2009) and to external plurals of nouns (cf. Hasselbach 2007). According to Hasselbach (2007) and other scholars, the nominal masculine plural -ū, the nominal feminine plural -ā and -āt (i.e., -ā+t) can be derived from the verbal system and the predicative adjective. As indicated in Hasselbach (2007), *-ā -ū* and -ī can be used for both external plurals and internal plurals of nouns and verbs.

Verbs (as in Amharic, Tigrinya and Tigre) and adjectives (as in Tigre and Amharic) have internal plural forms which look like the internal plurals of nouns in Tigre, Tigrinya and several other Semitic languages.

Buccellati (1996) and other scholars reveal that verbs, independent pronouns and pronominal suffixes in Semitic languages have *-a* (masculine) and *-i* (feminine) as primary gender markers; -ā (feminine), and -ū (masculine) as secondary gender markers, and *n* as a number marker.

In the verbs of EES languages, we can find different person, number, and gender exponents. As we can see below, however, they can synchronically be indicated by the same element.

Languages can have grammatical gender and natural (biological) gender. In this article, however, the discussion focuses on the former.

According to Baye (2009 E.C.: 120-121), Amharic nouns do not have a masculine grammatical gender marker, while *-it* (as in the case of *ayɨt'* 'mouse' and *ayɨt'itu* 'the mouse') can indicate the gender of feminine nouns. However, Baye also says the feminine marker *it* (1) occurs together with the definite article *-u* and (2) *it* has diminutive function. I assume *-it* appears to indicate smallness, diminutive etc. functions which can occur attached to the definite article.

Gender is considered an inherent quality of nouns. But in these languages, as in several other languages, gender inflections generally do not appear on nouns and EES nouns usually lack any gender specification (cf. also Getahun 1989 E.C.; Gut 1997; Hetzron 1997; Hudson 1997; Kogan 1997; Wagner 1997 among others).

The formal distinction between nouns and adjectives are not always clear (cf. Moscati *et al.* 1964 among others). Adjectives which make masculine/feminine gender distinctions (as in the case of *sərax'i* 'one who steals (m)' and *sərax'it* 'one who steals (f)') can be used as nouns. In the same way, I believe *säbɔay* 'man' was originally an adjective derived from *säbɔ* 'man' and *-ay* with the meaning 'belonging to'; while *säbäyti* 'woman' was originally an adjective derived from *säbɔ* 'man', -ay 'belonging to' and *-ti* (feminine marker). Some feminine human nouns may end in *-ti*. However, this *-ti* is unproductive feminine ending and can be observed only in rare cases (cf. Hudson 1997: 483 for the unproductive feminine ending *-t* in Amharic words like ɨnnat 'mother'). Generally speaking, EES nouns do not show gender distinctions. For instance, Wagner (1997: 492) says "Harari does not distinguish between genders through a form element". In Amharic and in Argobba, the gender of a noun is, as Hudson (1997: 464) puts it, "apparent in its choice of pronoun, agreement with the verb, determiners and the definite article suffix" (cf. also Getahun 1989 E.C.; Leslau 1995, 1997; Baye 2009 E.C.). Gender distinctions can be

observed as it affects the forms of the related words, a process called agreement. Nouns can be regarded as the "triggers", while other words may be the "targets" of changes. In the languages in question, these related words can be verbs, determiners (including pronouns), the number one, possessives, originally gerundive adverbs and adpositions.

In EES and other Semitic languages, person, number and gender can be marked by one element or by different elements. In the second person feminine plural (1a), for instance, we see the elements *k- ɨ -n-a* (<*kina*). The element *k* marks second person, ɨ (<*i*) marks primary feminine gender, *n* marks number; while *a* marks secondary feminine gender. In (1b-c), *-a* and *-u* mark third person feminine plural (3fpl) and third person masculine plural (3mpl) respectively. But in (1d-e), *-a* and *-u* mark feminine plural (fpl) and masculine plural (mpl) respectively. Consider the following:

(1) a. nägär-kɨna-ni
tell (perf.) -2fp-me
'you (2fpl) have told me',
b. nägär-a-ni
tell (perf.)-3fpl- me
'they (3fpl) have told -me'
c. nägär-u-ni
tell (perf.)-3mpl-me
'they (3mpl) have told me'
d. yɨ-nägr-a-ni näbär-a
3-tell (impf.)-fpl-me were-fpl
'they (f) were telling me'
e. yɨ-nägr-u-ni näbär-u
3-tell-(impf.)-mpl-me were-mpl
'they (m) were telling me'
Tigrinya

Amharic does not distinguish gender in the plural. Thus, the element *u* marks plural (pl) in (2a-b) and third person plural (3pl) in (2c) (cf. Moscati *et al.* 1964; Lipiński 1997 among others for Hebrew *-u* which marks plural in the perfective).

(2) a. näggär-ačč̣ih-u- ňň (ati+kum-+u + -ni > ačč̣čɨhuňň)
tell (perf.) -2pl-me
'you (2pl) told me'
b. yɨ-nägr-u-ňň näbbär
3-tell (impf.)-pl-me was
'they (pl) were telling me'
c. näggär-u-ňň
tell (perf.)-3pl-me
'they (3pl) told me'
Amharic

In Tigre, *-a* in (3a) and *-w* in (3b) indicate third person feminine plural (3fpl) and third person masculine plural (3mpl) respectively. In (3c-d), *-a* and *-o* (<*aw*) mark feminine plural (fpl) and masculine plural (mpl) respectively. Observe the following:

(3) a. qanṣa-y-a
get up (perf.)-3fpl
'They (f) got up'
b. qanṣa-w
get up (perf.)-3mpl
'They (m) got up'
c. tɨ-qanṣ-a
2-get up (impf.)-fpl
'you (fpl) are getting up'
d. tɨ-qanṣ-o
2-get up (impf.)-mpl
'you (mpl) are getting up'
Tigre

(Raz 1983:55-56)

In Gɨʔɨz, -ā in (4a) and -ū in (4b) indicate third person feminine plural and third person masculine plural respectively as shown in the following:

(4)	a. nagar-ā	b. nagar-ū	Gɨʔɨz
	speak(perf.)-3fpl	speak (perf.)-3mpl	
	ʿthey (f) spokeʾ	ʿthey (m) spokeʾ	

As indicated above, *-a* in (1a) marks feminine gender, while *-a* in (1d) marks feminine plural (cf. Buccellati 1996; Lipiński 1997 among others for similar cases in other Semitic languages). In the prefix conjugations, we find *y/t* to mark person, while *-a* and *-u* indicate feminine plural and masculine plural respectively in Proto-Semitic and in different Semitic languages including EES (1d-e, 2b). If there is a vowel *a/ä* before *u/w*, we can observe *a/ä+u/w=o* (cf. Leslau 1995; Tesfay 2002 among others). As in the case of *qäwämä=qomä* 'he stood up' (in Amharic and in Tigrinya), we can have *o* in *yɨfläṫṫ-o* 'let him know it' in Tigrinya that can be compared to Amharic -äw in *yɨsbär-äw* 'let him break it'. As in the case of other Semitic languages, *-a* and *-u* in the affirmative form of the imperative (imper.) EES verbs can indicate second person feminine plural and second person masculine plural respectively. If we compare the Tigrinya examples in (5ai, 5bi) and in (5aii, 5bii), however, we observe the second person marker *-t* appears on the surface in the negative forms of the verbs in the latter as in the following:

(5)	ai. nɨgär-a	aii. ʔay-ti-ngär-a	Tigrinya
	tell (imper.)-2fpl	do not you-tell-2fpl	
	ʿyou (2fpl) tellʾ	ʿyou (2fpl) do not tellʾ	
	bi. nɨgär-u	bii. ʔay-ti-ngär-u	
	tell (imper.)-2mpl	do not you-tell-2mpl	
	ʿyou (2mpl) tellʾ	ʿyou (2mpl) do not tellʾ	

If we compare the Amharic examples in (6a-b), we observe the second person marker *-t* appears on the surface in the negative form of the verb in (6b) as in the following:

(6)	a. nɨgär-u	b. ʔatti-ngär-u	Amharic
	tell (imper.)-mpl	do not you-tell-mpl	
	ʿyou (2pl) tellʾ	ʿyou (2pl) do not tellʾ	

As indicated above, the originally secondary gender marker can indicate both gender and number in the imperfective (impf.) form. In the Tigrinya verb *ti-wɨssɨn-u* 'you decide (2pl)', for instance, *t* indicates second person, while the originally secondary masculine gender marker *-u* indicates both masculine gender and number (plural). Hence, *-u* in *ti - u*, indicates both masculine and plural and we observe a syncretism of number with gender in the imperfective of Tigrinya (cf. Adger and Harbour 2007 for syncretism; Tesfay 2016 for syncretism in EES). Amharic does not distinguish gender in the plural. The originally masculine gender marker *-u* indicates both genders. In the imperfective third person plural of Amharic, we find an amalgam of number and gender. As in the case of (2b) above, we observe a syncretism of number with the originally secondary gender marker in the imperfective form of Amharic. In the perfective (perf.) forms such as *wässän-ačč̣ihu* 'you (have) decided (2pl)' or *näggär-ačč̣ihu-ňň* as in (2a), I assume *ačč̣ihu* is derived from *at+kumu* (<*kanu*) >*at+ḥuwu* >*ačč̣ihu* (cf. Tesfay 2016: 175-177 for more details).

In the above examples, we observe *-a* and *-u* that occur attached to verbs.

EES languages have also *-a* and *-u* (or their allomorphs) that occur attached to nouns, pronouns, determiners, gerundives (ger.), adpositions and adverbs with gerundive forms as illustrated in the examples in (7-9). The elements *-u* and *-a* occur attached to nouns as in (7a-b):

(7)	a. sɨm-u	b. sɨm-wa	Amharic
	name -3ms	name-3fs	
	'his (the) name'	'her (the) name'	

while in (8a-b), *-u* and *-a* are suffixed to pronouns as in the following:

(8)	a. hɨt-u	b. hɨt-a	Tigre
	hɨt-3ms	hɨt- 3fs	
	'he'	'she'	

(Raz 1997: 448)

Furthermore, *-u* and *-a* occur suffixed to determiners ((9a-b), gerundives (9c-d), adpositions (e-f), adverbs with gerundive forms (g-h)). Observe the following:

(9)	a. ʔɨt -u	b. ʔɨt -a	Tigrinya
	'the (3ms)'	'the (3fs)'	
	c. wässin-u	d. wässin-a	
	decide (ger.)-3ms	decide (ger.)-3fs	
	'he has decided'	'she has decided'	
	e. kab-ʔu	f. kab-ʔa	
	from-3ms	from -3fs	
	'from him/from there'	'from her/from there'	
	g. qäl'ɨf-u	h. qäl'ɨf-a	
	quick-3ms	quick-3ms	
	'quickly'	'quickly'	

In the examples in (7-9) and in (Tesfay 2016), we observe that gender can be marked by *-u* in the masculine and by *-a* in the feminine (or their allomorphs).

In the adjectives, however, gender can be indicated by *-a-, -u-, -i-, -ti* or *-it.* In Tigrinya adjectives which describe colour, size, weight, concentration, depth or height of nouns, we see the forms in (10a-b). The form *cä*(*a*)*c*(*c*)*ac* as in (10f-g) and the form *cä*(*a*)*cuc* >*cicuc* as in (10ci, 10ei) are participles. In (10a-b), we see *-a-* (feminine) and -i- (masculine) gender markers. In (10c, 10e), we observe *-u-* (masculine) and *-ti* (feminine) gender markers. But in (10f-g), *-a-* indicates both feminine and masculine genders. Moreover, Tigrinya has the active participle form as in (10di, 10dii). In (10dii), *-t*, indicates feminine gender. Consider the following:

(10)	ai. qäṫ'ṫ'an 'thin (f)'	aii. qäṫ'ṫ'in 'thin (m)',	Tigrinya
	bi. ħaṣṣar ' short (f)'	bii. ħaṣṣir 'short (m)',	
	ci. fiṫ'ur 'creature/created (m)'	cii. fiṫ'irti[2] 'creature/created (f)'.	

[2] The element *-i* following *-t* in words such as *fiṫ'irti* is an epenthesis.

di. qätali 'killer (m)' dii. qätalit 'killer (f)'
ei. sɨbur 'brocken (m)' eii. sɨbɨrti 'brocken (f)'
f. säbar 'brocken (m and f)' g. ħaffar 'shy (m and f)'

In the examples in (11ai-aii), *-a-* (feminine) and *-i-* (masculine) are gender markers. In (11bi-bii) too, *-u-* (masculine) and *-t* (feminine) are gender markers. Observe the following:

(11) ai. ħaddas 'new (f)' aii. ħaddis 'new (m)' GɨʕɨZ
bi. nɨgur 'spoken (m)' bii. nɨgɨrt 'new (f)'

Furthermore, the examples below show that *-a-* in (12ai), *-i-* in (12aii), *-u-* in (12bi) and *-at* (<*a*+*t*) in (12bii) are gender markers in Tigre which correspond to their counterparts in other EES languages. Consider the following:

(12) ai. ħadas 'new (f)' aii. ħadis 'new (m)' Tigre
bi. sɨbur 'broken (m)' bii. sɨbrat 'broken (f)'
(Raz 1983: 33-34)

In Amharic too, we find vowels which correspond to North Ethio-Eritrean Semitic *-a-, -i-,* and *-u-* in adjectives as in the following:

(13) a. qäċċɨn 'thin (m and f)' b. ʔaċċɨr 'short (m and f)' Amharic
c. räjjɨm 'tall (m and f)'
(14) a. t̓ɨqqur 'black (m and f)' b. sɨwwur 'hidden (m and f)'
c. fɨṣṣum 'complete (m and f)' d. kɨbur 'dear' (m and f)
(15) a. qällal 'simple (m and f)' b. käbbad 'heavy (m and f)'
c. fät̓t̓an 'fast (m and f)'

Tigrinya *qät̓t̓in* in (10aii) and ħaṣṣir in (10bii) correspond to (13a) and (13b) respectively. Besides, if we compare the adjective forms of Tigrinya in (10ci) and in (10ei) with those in (14a-d) in Amharic we observe that they have the same -ɨ -u vowel pattern. We also see that the vowel pattern of the adjective forms in (10ai and 10bi) correspond to those in (15a-c) in Amharic. The Amharic forms, however, do not distinguish gender. In fact, the Tigrinya forms in (10f-g), or Tigrinya adjectives like *däffar* 'courageous (m and f)' can, as in the case of Amharic, indicate both genders.

In (16), we observe that Amharic adjectives have internal plural forms which are similar to those of Tigre (different from those of Tigrinya) as shown below:

(16) ai. t̓ɨyyɨm 'dark brown (m and f)' aii. t̓ɨyayyɨm (m and f pl) Amharic
bi. aċċɨr 'short (m and f)' bii. aċaċċɨr (m and f pl)
ci. räjjɨm 'tall (m and f)' cii. räjajjɨm (m and f pl)
di. sɨbär 'you break' dii. säbabɨr 'you break repeatedly'
ei. säbbär-ä 'he broke' eii. säbabbär-ä 'he broke repeatedly'

Tigre adjectives have internal plural forms similar to those of Amharic (different from those of Tigrinya) as illustrated in (17) below:

(17) ai. ħaċir 'short (m)', aii. ħaċār (f), aiii. ħaċāyɨr (m and f pl) Tigre
bi. ħadis 'new (m)', bii. ħadās (f), biii. ħadāyɨs (m and f pl)
ci. fadāb 'brave (m)', cii. fadābit (f), ciii. fadāyɨb (m and f pl)
(*Ibidem*)

As indicated earlier, EES gender and number markers occur affixed to verbs, nominals, pronouns, determiners etc. In *kina* (1a), for instance, we observe that *k, i, n* and *a* mark second person, primary feminine gender, number and secondary feminine gender respectively. In (1b), however, *-a* marks 3fpl, while in (1d), *-a* marks feminine gender and plural. In (7-9), *-a* and *-u* mark feminine and masculine genders respectively. In (10-12), *-a-* and *-t* mark feminine gender, while *-u-* and *-i-* mark masculine gender. In (16-17), *-a-* is a marker of plurality.

Thus far, we can see that *a, u*, *i* and *-t* can indicate gender and/or number in EES languages. In (3), I will discuss these elements in the languages in question.

3. Relating Gender and Plural Markers

In Semitic languages, *u, i,* and *a* are gender markers. But the forms we use as gender markers may occur in the plural forms. According to Hasselbach (2007: 124-125), -āt (primarily associated with feminine) is the most common external plural, while nominative (masculine) and oblique (masculine) *-ū* and -ī which occur in a limited number of languages are second common external plurals. Greenberg (1955) identifies *a* as a marker of nominal and verbal plurality in Afroasiatic languages inserted in a consonant-vowel pattern. In fact, scholars assume that internal plurals of nouns and verbs are related (cf. Greenberg 1955; Benmamoun 2003 among others).

As indicated above, Hasselbach (2007) and other scholars assume that External and internal plurals in Semitic languages are related to the gender markers.

In Ethio-Eritrean Semitic languages too, we can observe that internal plurals of nouns and verbs are related among themselves and to the gender marking elements (cf. Tesfay 2009, 2016).

The examples in (1-9) illustrate that *a* and *u* can indicate feminine and masculine genders respectively in EES languages. Moreover, the examples in (10-12, 17) show that *-u-*, *-i-* and also *-a-*, *-t* (in *-ti* and in *-it*) can indicate masculine and feminine genders respectively. But these elements may also show number. Moreover, some of these elements may occur together to mark plurality. In Tigrinya, for instance, the suffixes *-ti*,[3] *-t* and the vowel *-a-* in an affix or within the stems indicate feminine gender, while (1) *a* extended by t as in *-at*, (2) *a…ti* >ä… *ti* (i.e. *a* within the stem suffixed by *-ti* >ä….*ti*), or (3) *a* >*ä*[4] followed by *w*[5]+(*t*) >äw(*t*) >*o*(*t*) mark plural (cf. Lowenstamm 1991 among others for *aw* >*oo* in Proto-Semitic).

In the adjectives of Tigrinya, Tigre and Gɨʕɨz, we find gender markers *-a, -u-, -i-, -t* (in *-it* and *-ti*) as shown in (10-12) above. In Amharic too, we find vowels which correspond to North Ethio-Eritrean Semitic *-a-, -i-,* and *-u-* in adjectives as indicated in (13-15). As we can observe

[3] The element *-ti* in *nigis-ti* 'queen' (in Tigrinya) corresponds to *-t* in *nigis-t* 'queen' (in Gɨʕɨz).
[4] *a* and ā in Tigre, Harari and Proto-Semitic correspond to ä and *a* in Tigrinya and Amharic.
[5] The element *w* is a vocalic equivalent of ū/u (cf. Hasselbach 2007: 128).

from the examples in (16-17), Amharic and Tigre adjectives have internal plurals (more or less similar to the nominal and verbal internal plurals of Tigrinya) different from their adjective counterparts in Tigrinya.

Besides, we see EES nominals whose plurals are expressed by attaching suffixes to the stems. These suffixes are *-očč, -an, -yan, -yat* in Amharic, -āt*, -otāt, -ot, -ač, -at,* and *-am* in Tigre and *-at, -tat, -an, -yan, -yat, -ti, -o,* and *-ot* in Tigrinya (cf. Raz 1983 for Tigre; Getahun 1987 E.C. for Amharic; Leslau 1995; Tesfay 2003 for Tigrinya; Baye 2009 E.C.). Tigre has also -āt and *-at* which indicate feminine plural and feminine singular respectively.

As we can observe from the discussion below, the plural forms are related among themselves and with the gender elements. The element *-a-/-ā-* as in (10ai, 17bii) indicates feminine gender, while *-a-/-ā-* as in *mänabɨr* 'chairs' (a plural form of *mänbär* 'chair') in Tigrinya and as in (17ciii) in Tigre marks plural. In (18a-b), *-a-* and *-i-* indicate feminine and masculine genders, while in the rest of the examples we observe a verbal and nominal plural marker *-a-* in the cvcacvc pattern in Tigrinya as in the case of *säbabär-ä* 'he broke repeatedly' (a plural form of *säbär-ä* 'he broke') as exemplified in (18c-g):

(18)	ai. qäṫṫan 'thin(f)'	aii. qäṫṫin 'thin(m)',		Tigrinya
	bi. ṣällam 'black(f)'	bii. ṣällim 'black(m)'		
	ci. mänṣäf 'carpet'	cii. mänaṣɨf 'carpets'		
	di. bäggiʕ 'sheep (sg)'	dii. ʔabagiʕ 'sheep(pl)'		
	ei. sɨbär 'you break'	eii. säbabɨr 'you break repeatedly'		
	fi. mänṫil-u 'he snatched'	fii. mänaṫil-u 'he snatched repeatedly'		
	gi. säbär-ä 'he broke'	gii. säbabär-ä 'he broke repeatedly'		

In Harari too, we find a verbal plural cvcācvc as in (19) below:

(19)	a. sabara 'he broke'	b. sibābara 'he broke repeatedly'	Harari
			(Wagner 1997: 494)

In Tigre, we see a feminine gender marker -ā- (20ai, 20bi) and a masculine gender marker *-i-* (20aii, 20bii) as in the following:

(20)	ai. ħaċċār 'short (f)',	aii. ħaċir (m),	aiii. ħaċāyɨr (m and f pl)	Tigre
	bi. ħaddās 'new(f)',	bii. ħaddis (m),	biii. ħaddāyɨs (m and f pl)	
				(Raz 1983: 33-34)

Furthermore, Tigre has a verbal and nominal plural marker *-ā-* inserted in the plural pattern cvcācvc as in *kanāfɨr* 'lips' (plural of *kanfar* 'lip') or in (20aiii, 20biii) above, and in (21aii, 21bii and 21cii) below:

(21)	ai. sabar-a 'he broke'	aii. sabābar-a 'he broke repeatedly'		Tigre
	bi. manṣaf 'carpet'	bii. manāṣɨf 'carpets'		
	ci. bɨgguʕ 'sheep (sg)'	cii. ʔabāgɨʕ 'sheep (pl)'		
				(1983: 19-20, 53)

The element -ā-/-a- inserted in the verbal and nominal internal plural cvcācvc/cvcacvc (cvcācvc >cvcacvc) pattern of Gᵻʕiz, Tigre and Tigrinya shows plurality. As we observe in (16), however, the pattern of the plural forms of verbs and adjectives can become cvcaccvc in Amharic as in the case of *aċċɨr* 'short (m and f sg)' and *aċaċċɨr* 'short (m and f pl)'. In Tigre too, we see the plural forms of adjectives *cacācɨc*/*caccācɨc*[6] as in (20aiii) and in (20biii).

As indicated above, the element -ā- in Tigre and Harari correspond to *-a-* in Tigrinya. In (18ai,18bi, 20ai, and 20bi) a/ā marks feminine gender, while in (18cii, 18dii, 18eii, 18fii, 18gii and in 20aiii, 20biii, 21aii, 21bii, and 21cii), a/ā marks plurality. The vowels *-a-* and *-i-* in Tigrinya ħaddas 'new(f)' and ħaddis 'new(m)' indicate feminine (f) and masculine (m) genders respectively. In Tigre too, we have *-ā* in ħadās 'new(f)' and *-i-* in ħadis 'new(m)' which indicate feminine and masculine genders respectively. In Tigre, the plural of ħadās and ħadis is ħadāyɨs 'new (pl)'. In Tigrinya, the plural of ħaddas and ħaddis is ħaddästi 'new (pl)'. In both Tigrinya and Tigre the vowel -ā or -ā >*-a* in ħad(*d*)*as*/ħad(*d*)ās 'new (f)' indicates feminine gender. In Tigre, the vowel -ā in ħadāyɨs 'new (m and f pl)' marks plural. In Tigrinya, the vowel *a* >*ä* and the suffix *-t* in ħaddästi 'new (m and f pl)' (*-i* in *-ti* is an epenthesis) mark plural. In Tigre too, feminine gender can be marked by *t* as illustrated in (22bi, 22di).

(22)	ai. gɨrrɨm 'handsome (msg)'	aii. gɨrrumām (mpl)	Tigre
	bi. gɨrrɨmɨt 'beautiful (fsg)'	bii. gɨrrumāt (fpl)	
	ci. qɨrub 'near, kin (msg)'	cii. qɨrubām (mpl)	
	di. qɨrbɨt 'near, kin (fsg)'	dii. qɨrubāt (fpl)	
	ei. sɨbur 'broken (msg)'	eii. sɨburām (mpl)	
	fi. sɨbrat 'broken (fsg)'	fii. sɨburāt (mpl)	

(1997: 449-450)

Moreover, feminine gender can be marked by *t* in Tigrinya as shown in (23aii, 23bii, 23cii, 23dii), while the plural can be marked by two feminine gender markers (*a*+*t*) as in (23aiii, 23biii, 23ciii and 23diii) below:

(23)	ai. mɨluʔ 'full (msg)'	aii. mɨlɨʔti 'full (fsg)'	Tigrinya
	aiii. mɨluʔat 'full (m and fpl)'		
	bi. ʕɨwwur 'blind (msg)'	bii. ʕɨwwɨrti (<ʕɨwwur+ti) 'blind (fsg)'	
	biii. ʕɨwwurat 'blind (m and f pl)'		
	ci. kɨbur 'dear (msg)'	cii. kɨbɨrti (<kɨbur+ti) 'dear (fsg)'	
	ciii. kɨburat 'dear (m and f pl)'		
	di. fɨluṭ 'known (msg)'	dii. fɨlɨṭti (> fɨluṭ+ti) 'known (fsg)'	
	diii. fɨluṭat 'known (m and f pl)'		

As in the case of Tigre and Tigrinya, *t* in Gᵻʕiz marks a feminine gender as in (24bi). The elements *-an* and *-at* are used as masculine plural (24aii) and feminine plural (24bii) gender markers.

(24)	ai. kɨbur 'dear (fsg)'	aii. kɨbur-an (mpl)	Gᵻʕiz
	bi. kɨbɨrt (fsg)	bii. kɨburat (fpl)	

[6] Short *-a-* and -ā- (long *-a-*) in Tigre cacācic/caccācic correspond to -ä- and *-a-* in Tigrinya cäcacic.

In (25a) and in (25c-f), we can see that the Amharic forms can be used for both genders. Thus, *-t* in (25b*)* appears to mark smallness, diminutive etc. Consider the following:

(25)	a. and	b. andit (< and-it)	Amharic
	one	one-f	
	'one(m)'	'one (f)'	
	c. and set	d. and wänd	
	'one woman'	'one man'	
	e. hulät set-očč	f. hulät wänd-očč	
	two woman-pl	two man-pl	
	'two women'	'two men'	

As illustrated in the examples above, the elements *a*, *t* or *a+t* can indicate feminine gender. But these feminine gender markers can also indicate plural. In Tigrinya, the two feminine gender markers *a+t*, can, as in (23aiii, 23biii, 23ciii, 23diii), indicate plural (m and f). In Tigre as in (22eii) and Gɨʕɨz as in (24bii), the feminine plural is marked by two feminine gender markers, ā/*a* + *t*. According to Hasselbach (2007), it is developed through the spread of /t/ from a feminine singular after it was interpreted as a general feminine marker. In (22cii, 22eii), we also observe that Tigre indicates number by ā+m in the masculine. As we can see in (24aii), Gɨʕɨz marks masculine plural by *a+n*. The morpheme *n* is an Afro-Asiatic number marker which can become *m* in the masculine as in the case of Tigre or Berber (cf. Siddiqi 2009 for examples from Berber; Tesfay 2016 for examples from Tigrinya).

In pronouns, the number element *n* (that can be realized as *m* in the masculine) can mark number (cf. Buccellati 1996; Lipiński 1997; Siddiqi 2009; Tesfay 2016) in EES and other Semitic languages. The element *-n/m* can also mark number in EES possessive suffixes. In verbs too, number can be indicated by the elements *n/m*, *-a* or *-u.* In EES Nouns and adjectives, *-at, -ti, -ač, -oč, -o/ot* can indicate plural. We observe palatalization in EES languages. According to Lowenstamm (1991), *au* becomes *ō* in Proto-Semitic. I assume the suffixes *-ač* in Tigre and *-oč* in Amharic are derived from *-a+ti* (i.e., *-a +ti* > *-ač*) and *-o+ti* (i.e., *-o + ti* > *-oč*) respectively.

As we can see in the examples above, the elements *t* (realized as *-t, -ti* or *it*) and *a* can mark feminine gender. However, these elements can also be used as plural markers.

In Tigre, -āt can mark plural (cf. Raz 1983: 17-18), while diminutives and paucatives can be marked by *-at* or *-it* and by -āt respectively (1983: 25-26).

In the languages in question, we see syncretism that can be defined as the representation of different combinations of morphosyntactic values by the same form. In the third person plural for instance, Tigrinya syncretizes number and person with gender in the perfective. Thus, in (1b) and (1c) we see -a (3fpl) and -u (3mpl) respectively. However, number can also be marked by *au/aw* > *o* / -äu/äw > *o* and *-aut/awt* > *ot* / -äut/äwt > *ot* (short *a* becomes ä in several EES languages) and their allomorphs.

As discussed above, gender and number markers are related. Hence, (1) the feminine gender markers *a/ā,* and *at*/āt can function as internal or external plural markers, (2) the masculine gender marker *u* can function as a plural marker. (3) *au*/äu > *o* or *aut*/äut > *ot* can become external plurals (4) *a* + *ti* > *-ač* and *au+t*/äu+ti > *-oč* can be used as external plurals.

4. Number and Gender as classifiers

We have said earlier that gender is not overtly marked on EES head nouns. In fact, Lipiński (1997: 233) says "the formal distinction between masculine and feminine is not an original feature of Semitic languages, as shown by the many basic feminine nouns without any special morpheme". Lipiński notes that "This opinion is apparently confirmed by the South Ethiopic idioms of the Gurage group which have no feminine mark" (*ibidem*). However, nouns derived from earlier adjectives and participles can show gender distinctions as illustrated in (26-28) Tigrinya examples below:

(26)	a. wäladi	b. wäladit	c. wälädi (< wäläd-ti)	Tigrinya
	parent-m	parent-f	parent-pl	
	'parent (m)'	'parent (f)'	'parents'	
(27)	a. ʔamani	b. ʔamanit	c. ʔamän-ti	
	believer-m	believer-f	believer-pl	
	'believer (m)'	'believer (f)'	'believers'	
(28)	a. mɨrux	b. mɨrɨx-ti	c. mɨrux-at	
	prisoner (m)	prisoner (f)	prisoner-pl	
	'prisoner (m)'	'prisoner (f)'	'prisoners'	

The word ʔanɨsti (29b) can be related to Proto-Semitic, *nš* 'man (kind)' and *ʔnθ* 'woman'. The early function of *-t* in *ʔanɨsti* may be a classier, a feminine gender marker or a plural marker. For the current speaker of the language, however, *-t* in *säbäyti* (29a) and in *ʔanɨsti* (29b) does not play such a role. The element *-t* is only regarded as a part of the word *ʔanɨsti* or *säbäyti* and not as a morpheme *-t*. Consider the following:

(29)	a. säbäyti	b. ʔanɨsti	Tigrinya
	'woman'	'women'	

But *-t* in (26b, 27b), and *t-* in (28b) indicate feminine gender, while the vowel *a >ä* (in the cv pattern of the stem) followed by *t-* in (26c, 27c) and *-at* in (28c) are plural markers.

In Tigre, *-at/-it*[7] and -āt can indicate feminine singular (fsg) and feminine plural (fpl) respectively as in the following:

(30)	a. nāfɨʕ	b. nāfʕat	c. nafʕāt	Tigre
	'useful (m)'	'useful (fsg)'	'useful (fpl)'	
(31)	a. qadām	b. qadāmit	c. qadāmyāt	
	'former'	'former (fsg)'	'former (fpl)'	

(Raz 1983: 33)

Languages like Burmese never use a simple numeral, as "one man", but employ auxiliaries, affixes, words or items which signify the class to which the name belongs, the use to which it is put, the resemblance in shape, or form of the referent etc. In EES languages, as in many other world languages, we do not expect to see such classifiers.

[7] The element *i* in *it* is an epenthesis (cf. Raz 1983: 7 for typical Tigre sequence composed of cv and cvc).

However, EES languages, as in the case of European languages, use measure words, which are required for counting mass nouns. Measure words denote a particular measurement of something (e.g. a drop, a pint, a cupful). They can be used to quantify mass nouns that indicate things without inherent countable units as in (32-33). In this sense, measure words are also known as mass classifiers. Observe the following examples from Amharic:

(32) a. hulät bɨrćiqo təĵ Amharic
two glass local wine
'two glasses of local wine'
b. assɨr doniyya ťef
ten sack ťef
'Ten sacks of ťef'

In the same way, we have similar Tigrinya examples in (33)

(33) a. ħadä ʕɨtro may Tigrinya
one jar water
'a jar of water'
b. sälästä ṣiʕnät gämäl ćäw
three load camel salt
'three camel load of salt'
c. sälästä ʔɨnfɨx'ti mäšäla
three ʔɨnfɨx'ti sorghum
'three ʔɨnfɨx'ti of sorghum' (ʔɨnfɨx'ti = about 20 kilos)
d. ħadä saṣun kɨdawɨnti
one box clothes
'one box full of clothes'

In general, we have said earlier that gender is not overtly marked on EES head nouns (it is not marked by a gender marker attached to the noun). In Amharic for instance, Leslau (1995: 161) says distinction in gender is indicated by the gender of the definite article, by gender of the verb etc. If we compare the Amharic phrases *and set* 'one woman' (25c), *andit set* 'one woman' (25b) and *and wänd* 'one man' (25d), the element *-it* does not appear to show gender. It may have diminutive or classifying functions. If we take other Amharic examples, the element *-it* as in *ɨj* 'child, girl, boy' and *ɨj-itu* 'the girl', *bäg* 'sheep' and *bägitu* 'the ewe' which occurs attached to a definite article may have similar functions. Moreover, Some EES languages have the same form for the singulars and plurals. In Muher, for instance, *fàrὰz* 'horse' can also mean 'horses'. In Tigre, the suffix *-at* (or *-tat*[8]) can be used to derive the singulative from the collective as in *rɨšaš* 'lead' and *rɨšaš -at* 'a bullet, piece of lead', *wagre* 'olive trees, fruits' and *wagretat* 'an olive tree, a piece of olive wood'.

Genders can be defined as classes of nouns reflected in the behaviour of associated words (cf. Corbett 1991: 3; Wälchli and Di Garbo 2019). If one says a language has 3 genders, it implies that there are 3 classes of nouns that can be distinguished syntactically. In fact, Corbett (1991: 135) argues gender systems arise from the use of nouns with classificatory possibilities.

[8] As *-tat,* as in *wagre-tat,* occurs suffixed to stems ending in vowels, the first *t* in *-tat* is inserted to break the impermissible sequence of the vowel *a* in *-at* and the preceding vowel.

Hasselbach (2007: 131-133) says -ā/-āt, *-ū* and -ī seem to be gender dependent associated with plurality used for both internal and external plurals. According to Fischer (1997: 193-194), the most common of the Arabic feminine markers is *-at*, (*-a*: and *-aːʔ* are mainly reserved for adjectives). Fischer argues the basic function of *-at* is to denote the particular distinguishing it from the general. According to Fischer, *-at* marks (i) the female in contrast to the generic term (which includes the male gender), (ii) the outstanding person in contrast to the common (the evaluative function may be included here) (iii) the singulative in contrast to the collective. Lipiński (1997), Buccellati (1996), Siddiqi (2009) and several other scholars show that the element *n* (which can become *m* in the masculine) is an Afro-Asiatic number marker. However, the data from Semitic languages reveal that number is frequently marked by gender elements *a, u* and *t*. Moreover, number can be marked by the combination of these gender markers. I assume EES gender and/or number markers (*a/ā, t, u/ū, a+t/*ā+t=*at/*āt, *a+u/ä+u=o, a/*ä+u+t=*ot, a+ti=ač, a+u+ti=oč*) can function as classifiers. However, further research is also needed.

5. Multi-Plurals

In this section, internal plurals followed by one or two external plural suffixes are referred to as multi-plurals. As indicated above, the languages in question can have internal plurals with the pattern $c^1vc^2ac^3vc^4/c^1vc^2āc^3vc^4$. The vowel *a/ā* following c^2 is a plural marker. The vowels after c^1 and after c^3 are usually ä/ā and ɨ respectively. In languages such as Tigrinya, for instance, we see forms like *sur* 'root' and *särawɨr/särawur* 'roots'. In languages like Tigre too, the vowels after c^1 and after c^3 are usually realized as *a* and ɨ respectively as in the case of *manṣaf* 'carpet' and *manāṣif* 'carpets'. But it is possible that the vowels after c^1 and after c^3 in the cv plural pattern can be similar to the vowels of the first and last syllables in the singular (cf. McCarthy 1982 among others for Arabic). If, for instance, we compare the words ħiṣan 'baby' and $ħ^1iṣ^2aw^3in^4ti$ 'babies' in Tigrinya, *liq* 'scholar' and $l^1iq^2aw^3in^4t$ 'scholars' in Gɨʕɨz and in Amharic, we observe that the vowels after c^1 in the internal plural patterns are similar to the vowels in the first syllables of their singular forms. Furthermore, the vowel following c^1 can also be realized as *a* if c^1 is a pharyngeal or a laryngeal as in (37b). As indicated above, the vowel between c^3 and c^4 in $c^1vc^2ac^3vc^4$ /$c^1vc^2āc^3vc^4$ is usually ɨ. However, it may be similar to the vowel in the last syllable of the singular as in the case of Tigrinya ʿamil 'client' and ʿamawil 'clients', *mändil* 'handkerchief' and *mänadil* 'handkerchiefs'. If the last syllable of the plural ends in ɨw, we observe ɨw>*u* as in Tigrinya *ʕataɨw>ʕataru 'jars' (plural of ʕɨtro 'jar'). If c^3 is geminated, the vowel *a* following c^2 can be realized as ä as in the case of ʕanjal 'fool' which forms its plural as ʕanajɨl in standard Tigrinya and ʕanäjjɨl in a dialect of Tigrinya. According to Lowenstamm (1991) (verbal communication), Ratcliff (1998) and other scholars, the change of *a* to ä following c^2 is substituted by the gemination of the following consonant (i.e. c^3). As mentioned above, the vowel v following c^3 can be similar to the vowel of the singular in the last syllable or can be realized as ɨ, *u* or *a*. If the vowel following c^3 is realized as *u*, the plural vowel *a* following c^2 can optionally be deleted as in the case of *ʔabaʕur/ʔabʕur* 'oxen' in Tigrinya. If the vowel following c^3 is realized as *a*, the plural vowel *a* following c^2 can be deleted as in the case of **ʔaħawat* >*ʔaħwat* 'brothers/sisters' (40b) which can be related to Arabic *ʔaḫawāt-un* 'sisters' (cf. Hasselbach 2007: 125) and to Arabic *ʔaḫawāni* 'brothers'. Moreover, *-at* or *-ti* can idiosyncratically be added to the internal plural pattern as in the case of *mäṣaħɨfti* 'books', *ʔataxɨlti/ʔatkɨlti* 'plants' (38b, 38c) or *ʔaxlabat* 'dogs' (36d). As in other Semitic languages (cf. Moscati *et al*. 1964 among others), the internal plurals of Tigrinya, Amharic, Gɨʕɨz and Tigre followed by *-at* or *-ti* suffixes can etymologically be regarded as double plurals. However, we

may find Amharic internal plurals followed by two external suffixes as in (41c, 42c). In this article, the multi-plurals and double plurals in (34b, 34c, 35c, 36d, 37d, 38b, 38c, 41b, 41c, 42b, 42c) can be referred to as multi-plurals.

The word in (34b) is etymologically a double plural form of the singular indicated in (34a). In (35b), the word *mämhɨran* (derived from *mämhɨr+an*) is a plural form of the singular word in (35a). Moreover, the words in (34c and 35c) are the multi plurals of the singulars in (34a and 35a) respectively as in the following:

(34)	a. liq	b. liqawɨnt	c. liqawɨntočč	Amharic
	'scholar'	'scholars'	'scholars'	
(35)	a. mämhɨr	b. mämhɨran	c. mämhɨranočč	
	'teacher'	'teachers'	'teachers'	

The forms in (36b-d), (37b-d), (38b-c), (39b) and (40b) are the plural forms of (36a), (37a), (38a), (39a), and (40a) respectively. (36b) and (37b) are the basic internal plural forms, while in (38b) *-ti* is added to the basic internal plural. In (36c-d, 37c-d, 38c, 39b and 40b), the plural vowel *a* following c^2 in the cv plural pattern is deleted. This may happen if the vowel following c^3 in the prosodic template of the plural (or cv plural pattern) is realized as *a* as in (36c, 37c, 39b, 40b) or *u* as in *ʔabaʿur/ʔabʿur* 'oxen' and/or if the cv pattern is followed by *-at* or *-it* as in (36d, 37d, 38c). In (36d), *-at* is added to the form in (36c). In (37d) too, *-ti* is, I assume, added to the form in (37b). The plural form of ħaw 'brother' in (39a) and ħawti 'sister' in (40a) is *ʔaħwat* in (39b and 40b) which is derived from the basic internal plural form **ʔaħawat* (similar to Arabic *ʔaħawat*). Consider the following:

(36)	a. kälbi	b. ʔaxalɨb	c. ʔaxlab (< ʔaxalab)	Tigrinya
	'dog'	'dogs'	'dogs'	
	d. ʔaxlab-at			
	'dogs'			
(37)	a. bäx'li	b. ʔabax'ɨl	c. ʔabqal (< ʔabaqal)	
	'mule'	'mules'	'mules'	
	d. ʔabqɨlti			
	'mules'			
(38)	a. täxli	b. ʔataxɨlti	c. ʔatkɨlti (<ʔataxɨl-ti)	
	'plant'	'plants'	'plants'	
(39)	a. ħaw	b. ʔaħwat (< *ʔaħawat)		
	'brother'	'brothers/sisters'		
(40)	a. ħawti	b. ʔaħwat (< *ʔaħawat)		
	'brother'	'brothers/sisters'		

The forms in (41b-c) are the plurals of (41a). The forms in (42b-c) are the plurals of (42a). However, there are formal differences between the forms in (41b and 42b) and those in (41c and 42c). Those in (41b and 42b) are etymologically double plurals, while in (41c and 42c) we find the Amharic external plural *-očč* suffixed to the double plurals as illustrated below:

(41)	a. mäṣħaf	b. mäṣaħɨft	c. mäṣaħɨftočč	Amharic
	'book'	'books'	'books'	
(42)	a. liq	b. liqawɨnt	c. liqawɨntočč	
	'intellectual'	'intellectuals'	'intellectuals'	

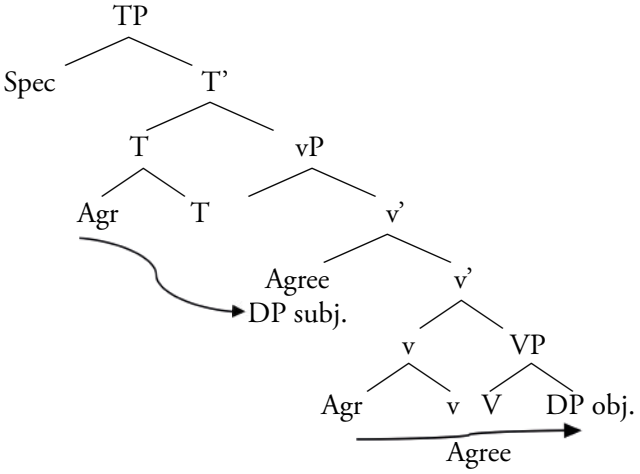
There are Amharic native speakers who think that the plurals of *mäṣhaf* and *liq* must be *mäṣhafočč* or *mäṣahift* (and not *mäṣahiftočč*) for the former, *liqočč* and *liqawint* (and not *liqawintočč*) for the latter. The plural of the English word *child* is *children*. But in the literature, we see that the plural of the word *child* was originally *childer* and *children* was considered a double plural. In Tigrinya, native speakers are not aware that the suffixes *-at* in (36d) and *-it* in (38b) are double plural markers. In fact, they are (ignoring historical antecedents) regarded as only part of the plural form. In Amharic too, the originally double plurals *liqawint/liqočč* and *mäṣahift/mäṣhafočč* (together with *liqawintočč* and *mäṣahiftočč*) may be accepted as plurals of their singular forms.

6. Position of Agreement morphemes on the Structure

In the literature, it is indicated that the structure that contains the nominative subject is a CP and the C selects the T, whereas the structure that contains the genitive subject is without CP. Thus, it is assumed to have a defective T. Scholars argue that the D that takes the defective T is allowed to license the genitive case (cf. Miyagawa 2012: 8, 126, 131, 134, 146).

Fuß (2005) and other scholars believe that Agr-nodes do not head their own projections in the syntax. Thus, Agr-morphemes can attach parasitically to other "substantial" functional categories (such as T). In Corbett (1991: 18), we see that gender is located at n, while number can be located at n, n and num or just num. According to Fuß (2005: 81-84), subject and object agreement morphemes occur attached to T and v respectively. The head complex [v Agr (v)] can enter into an Agree relation with the feature set of the object DP, while the head complex [T Agr (T)] initiates an Agree operation which targets the closest active DP with an appropriate feature content which leads to subject agreement and nominative licensing. Object and subject agreements are checked after v and T have been merged with their complements. In the latter T merges with vP, while in the former, v merges with its complement VP which contains the object as we can see from the tree below adapted from Fuß (2005: 83-84):

(43)



(*Ibidem*)

In the case of EES languages, I assume subject, object and genitive subject agreement morphemes occur attached to T, v and D respectively. In the latter case, I assume agreement is checked after D has been merged with its complement nP. I think the head complex [D Agr

(D)] may enter into an Agree relation with the feature set of the genitive DP.[9] Following recent literature, I assume the following:

a. Agents within DP are generated in the specifier of a little n whose complement is NP and the head of NP raises into n (cf. Adger 2003: 267-275, Fuß 2005 among others).

b. Roots are considered acategorial in that their syntactic category is contextually specified by combining with category-defining functional heads such as v, n, a (cf. Sato 2010: 16-19, Arad 2005: 42-47 among others).

c. APs branch from nP (cf. Adger among others).

EES languages have nouns which can correspond to Hebrew mišqalic andNon-mišqalic nouns. In the former, consonantal roots are combined with nominal patterns, while in the latter, nouns are made of syllabic roots and many of them can be borrowed (cf. Arad 2005). Thus, items like *därho* 'hen', *lam* 'cow' and *bet* 'house' are roots.

According to Arad (2005: 42), binyanim have vowel slots, but lack the vowels themselves, while mišqalim have their inherent vowels specified. According to Siddiqi (2009: 51-54), irregular nouns and irregular verbs of English are realized by one VI. In *John ate*, for instance, √eat and [PAST] are realized by one VI, while in *John killed,* √kill and [PAST] are realized by two VIs*, kill* and *-ed.* Taking such views into account, we can have Tigrinya plural forms such as *m¹än²ab³ir⁴* 'chairs' (a plural form of *m¹*än²b³är⁴ 'chair') that can be realized as one VI. Regarding the plurals like *säb-at* 'persons', however, the root and the suffixed plurals are realized by two VIs. In the internal plurals like *mälaʔixti* 'angels', *ʔataxilti/ʔatkilti* 'plants', and *ʔaxlabat* (<*ʔaxalabat) 'dogs', *-ti* and *-at* may be added idiosyncratically to the already plural and form a double plural (cf. also the examples in 34-42 above). The plural marking *a/ā* of the template may be deleted (McCarthy 1982 among others).

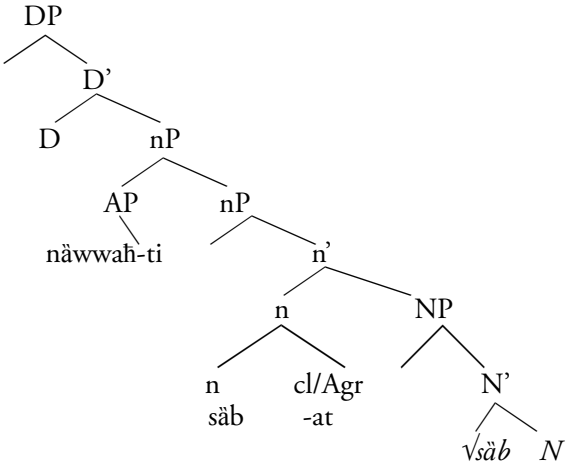
Taking the above indicated views into consideration, we can have Tigrinya examples *näwaħ-ti säb-at* (44b and 44d), *näwwaħ märaħit hagär* (44e) and *näwwaħ-ti märaħti hagär* (44f) which correspond to the structures in (45), (46) and in (47) respectively.

(44)	a. näwwiħ säb	b. näwwaħ-ti säb-at	Tigrinya
	tall (m) person	tall- pl person-pl	
	'tall (masculine) person'	'tall persons'	
	c. näwwaħ säb	d. näwwaħ-ti säb-at	
	tall (f) person	tall- pl person-pl	
	'tall (feminine) person'	'tall persons'	
	e. näwwaħ märaħ-it hagär	f. näwwaħ-ti märaħ-ti hagär	
	tall (f) leader-f country	tall -pl leader-pl country	
	'a tall (f) leader (f) of a country'	'tall (pl) leaders of a country'	

In (45), the root *säb* is inserted under the root node and merges with the category-defining head N. The root *säb* which occurs as a sister to N raises to n, while *näwwaħ* followed by the plural suffix *-ti* occurs in AP which branches from nP as in the following:

[9] Subjects in both nouns and clauses are generated within the projection of the lexical categories: N in the former and V in the latter (cf. Adger 2003; Fukui 2006 among others).

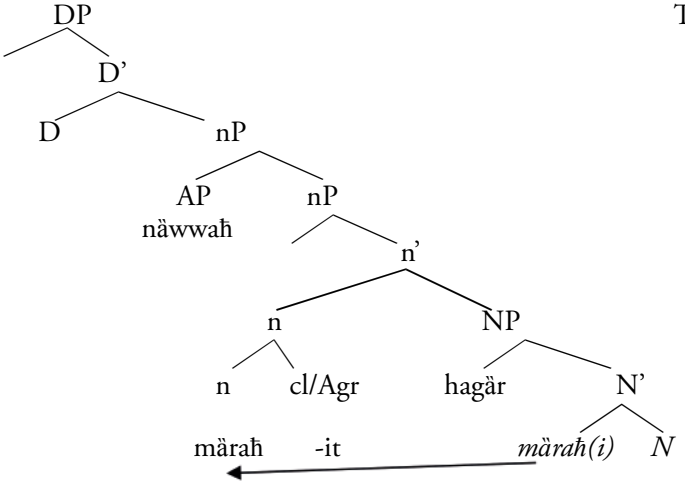
(45) Tigrinya



According to Adger (2003) and others, we can have nP (which corresponds to vP) between NP and DP, while APs can branch from nP. The head complex [n Agr [n]] may enter into Agree relation with the feature set of the DP under NP. In (45), I assume the plural element *-at* is added to n. I assume the plural suffixes *-at* and *-ti* are suffixed to *säb* and to *näwwaħ* respectively.

As we can observe below, *näwwaħ märaħ-it hagär* (44e) corresponds to the structure in (46). We can see that in (46), the head *märaħ*(*i*) raises into n, just as the head of VP raises into v (cf. Adger 2003: 275), while *hagär* occurs in the spec of NP. I assume the feminine marker *-it* occurs under n. In the adjective *näwwaħ* (which branches from nP) the vowel *a* marks feminine gender. I assume the actual ϕ-features of the Goal are copied onto the Probe in the initial postsyntactic module (after syntactic operations are complete) at MS, though the details may need further investigation. Let's see the following:

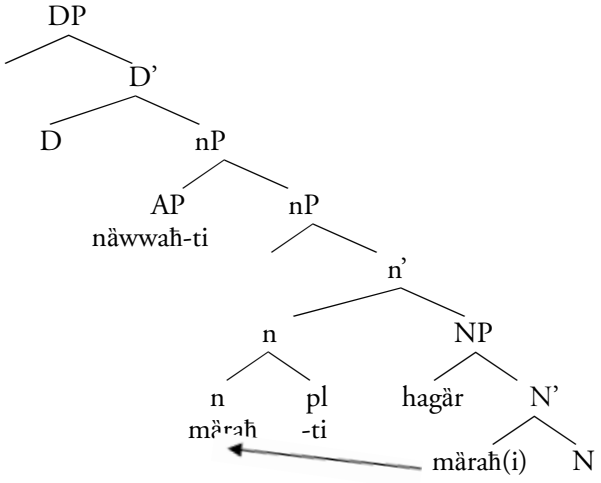
(46) Tigrinya



In (46), the feminine marker *a* in *näwwaħ* can be compared to the masculine marker *i* in *näwwiħ* 'tall (m)'.

We see that näwwaħti märaħ-ti hagär (44f) has a structure in (47). The head raises into *n* and gets the plural suffix *-ti*. The adjectival phrase (AP), which includes participles and adjectives of the language in question, branches from nP (cf. Adger 2003: 274-276).

(47) Tigrinya



As indicated above, Hebrew binyanim have vowel slots, but lack the vowels themselves, while mišqalic nouns have their inherent vowels specified. However, non-mišqalic nouns are made of syllabic roots. Mišqalic nouns are made of consonantal roots, while non-mišqalic nouns are made of syllabic roots (cf. also Arad 2005: 34-42). Taking such views into account, we can have the root *liq* in Amharic, Tigrinya and Gɨʕɨz. As in the case of *m*^{1}änb^{2}är^{3} 'chair', *m*^{1}än^{2}*ab*^{3}ɨr^{4} 'chairs', *saʔni* 'shoe' ʔ^{1}*as*^{2}*a*ʔ^{3}ɨn^{4} 'shoes', **l*^{1}*iq*^{2}*aw*^{3}ɨn^{4} is a possible plural of *liq* 'scholar'. However, the actual plural of *liq* is *l*^{1}*iq*^{2}*aw*^{3}ɨn^{4} +*-ti* 'scholars' in Tigrinya. Recall the noun liq and its plural form in Tigrinya (repeated here as 48b).

(48) Tigrinya

a. liq	b. l^{1}iq ^{2}aw^{3}ɨn^{4}-ti
'scholar'	'scholars'

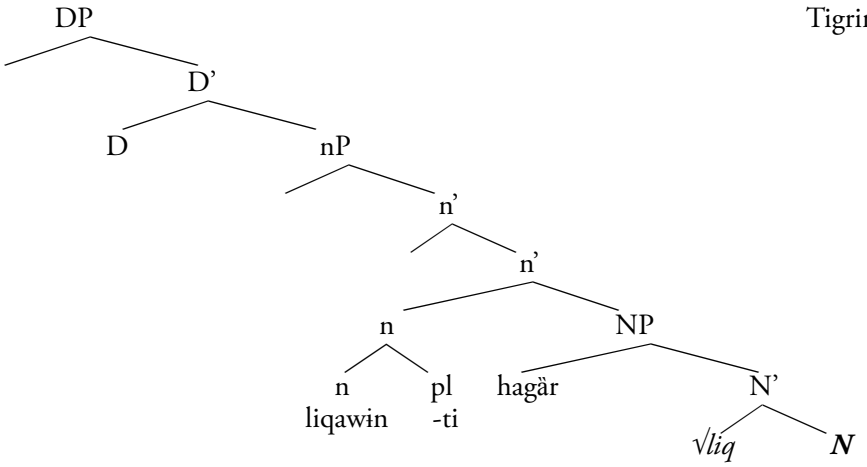
In Amharic too, **l*^{1}*iq*^{2}*aw*^{3}ɨn^{4} is a possible plural of liq, while the actual plurals are *l*^{1}*iq*^{2}*a-w*^{3}ɨn^{4} +*-t* 'scholars' or *l*^{1}*iq*^{2}*aw*^{3}ɨn^{4}+*-t*+*-očč* 'scholars'. Recall the Amharic plurals for liq (repeated here as 49b, 49d).

(49) Amharic

a. liq	b. l^{1}iq^{2}aw^{3}ɨn^{4}-t
'scholar'	'scholars'
c. liq	d. l^{1}iq^{2}aw^{3}n^{4}t-očč
'scholar'	'scholars'

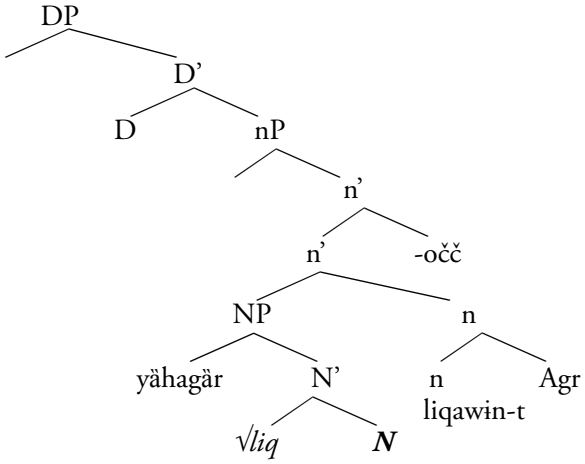
In the examples in (48b) and in the tree structure in (50), we see a double plural in Tigrinya. As illustrated in (50), Tigrinya *liq* forms a possible plural liqawɨn in n. The possible plural and the suffix *-ti* form a double plural. The double plural is the actual plural of Tigrinya *liq* as shown in (50).

(50) Tigrinya



In Amharic, we see double and multi plurals as in (51). In (51), Amharic *liq* forms a possible plural *liqawɨn* in n. The possible plural and the suffix *-t* form a double plural. The double plural is the actual plural of Amharic liq. Furthermore, Amharic can form the multi plural *liqawɨn-t-očč* by adding *-očč* to the double plural *liqawɨnt.*

(51) Amharic



In Amharic, we observe the suffixes on the right side of the tree structure. Compound words in Amharic are right headed and the suffixes indicating number occur attached to the head as in the case of *yämängɨst särratäňa* 'public servant' and *yämängɨst särratäňočč* 'public servants'. In languages like Tigrinya, compound words such as *särahtäyna mängɨsti* 'public servant' and *särahtäyna-tat mängɨsti* 'public servants' are left headed and the suffixes indicating number occur attached to the head. I have no intention to go into details. According to Arregi and Nevins (2012), however, the syntactic computation does not contain statements of linear order only of sisterhood and dominance. Fuß (2005) also assumes linear order is not completely determined by syntax and certain points regarding the inflectional markers are resolved in the post-syntactic morphological component.

7. Conclusion

In this article, I tried to give an overview on the relationship of some aspects of Semitic languages. More specifically, different EES gender and number markers are discussed.

The first observation made is that the vowel -ā which can be identified as a marker of nominal and verbal plurality in Afroasiatic languages survives in the internal plurals, external plurals, verbal plurals and nominal plurals of Semitic (including EES) languages. Moreover, the element *n* also marks number in Semitic and Afroasiatic languages.

The second observation is that Semitic languages employ the elements ā, ū, ī and *t* to mark gender. However, the elements ā, ū, ī, *-an* (<*a+n*) and *-at* (<*a+t*) also mark plural (cf. also Hasselbach 2007 for Ancient Egyptian, Middle Egyptian and Berber *-au* and *-aw* to indicate masculine plural). Thus, we see the interdependence of gender and number markers in nouns, adjectives and verbs of Semitic languages. The fact that the Semitic verbal system is based on an originally adjectival pattern[10] may be the reason why gender and number are marked by the same elements in nominal and in verbal forms (cf. Moscati *et al.* 1964, Lipiński 1997; Hasselbach 2007; Carver 2016 among others for the use of similar gender/number markers in the verbal system and on the predicative adjective and for the frequent lack of clear distinctions between nouns and adjectives in Semitic languages).

The third observation is that EES gender and number markers are similar or strikingly related to Semitic gender and number markers. In EES languages, ā or ā >*a*, and ū or ū >*u* can mark gender in verbs and in pronouns. But these elements can also mark gender and number or gender, number and person. Furthermore, EES languages employ ā or ā>a, *u*, *i* and *t* to mark gender in adjectives and participles. The gender markers ā/a, āt/*at* and *a..t*>*ä--t* can also be used as internal and/or external plural markers in nouns and/or in adjectives of EES languages. As in other Semitic languages, EES languages use ā or ā >*a*, ū or ū >*u,* āt/*at, n* and *an* to indicate number. However, plurality in EES is also indicated by the reflexes of these Semitic plural forms. Thus, EES languages have -ā/a, -āt/*at, -ačč* (<*-ati*), *-ū/u, -očč* (<-ä*u+ti*), *-o* (<-äw), *-ot* (<-äwt), *-n/m*, *-an*, *-yan*[11], *-yat*, ʔn[12] as plural markers. As in other Semitic and Afroasiatic languages, number can be marked by n/m. As in the case of other Semitic languages, gender markers alone or together with other gender or number markers can be used as number markers. Hence, we observe the interdependence of gender and number markers in EES languages too.

The fourth observation is that genders can be defined as classes of nouns reflected in the behaviour of associated words (cf. Corbett 1991 among others). Several previous works on Semitic languages show that the long vowels in verbs, adjectives and participles mark gender. Gender morphemes can correspond to classifiers. As indicated in Wälchli and Di Garbo (2019), I believe classifiers are similar to gender in that they function as classes of referents in the languages in question (cf. Corbett 1991; Kihm 2008; Manzini, Savoia and Tesfay 2018 among others).

It seems possible to assume that -ā or ā >a and -ū or *ū* >*u* were early gender morphemes in verbs and in adjectives (cf. Lipiński 1997: 341, 360 for the relationship between stative forms, verbs and nouns, Hasselbach 2007:132 for the originality of *-ū* and -ā on the verbal system

[10] According to Lipiński (1997: 336-337, 360), the aspectual category of the verbal system is based on the adjectival $c^1a\ c^2c^3$ pattern which can be represented by the Assyro-Babylonian verbal adjective. It developed to the stative/permansive form that became a perfective which can function as a verb. The stative, however, essentially represents the conjugation of a noun or an adjective.

[11] *y* in *-yan* and *-yat* is an epenthesis.

[12] The element ʔn can be related to the demonstrative hn and to the number element *n* (cf. Buccellati 1996; Lipiński 1997 among others).

and on the predicative adjective). I assume that *-ā* or *ā* >*a,* and -ū or *ū* >*u* (which may be followed by *-t, -u,* or *-n/m*) became markers of a section of a group as a subset of another group depending on gender or number. As in other Semitic languages, there is interdependence of gender with number and the use of the former on verbal and nominal forms may be regarded as more original than the latter. Nonetheless, I also assume further research is needed to explore the diachronic relationship between them.

References

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. New York, NY: Oxford UP.

Adger, David, and Daniel Harbour. 2007. "Syntax and Syncretism of the Person Case Constraint." *Syntax* 10 (1): 2-37.

Arad, Maya. 2005. *Root and Pattern: Hebrew Morpho-Syntax*. Dordrecht: Springer.

Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. New York, NY-London: Springer.

Baye, Yimam. 2016/2017 [2009]. *Yamariňňa Säwasiw [Amharic Grammar][,3rd edn.]*. Addis Abeba: Addis Abeba University Printing Press.

Benmamoun, Elabbas. 2003. "Reciprocals as Plurals." In *Research in Afro-Asiatic Grammar II: Selected Papers from the Fifth Conference on Afroasiatic Languages*, Paris 2000, ed. by Jacqueline Lecarme, 53-62. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company.

Buccellati, Giorgio. 1996. *A Structural Grammar of Babylonian*. Wiesbaden: Harrassowitz.

Buccellati, Giorgio. 1997. "Akkadian." In *The Semitic Languages*, ed. by Robert Hetzron, 69-99. London-New York, NY: Routledge.

Carver, Daniel E. 2016. "The Akkadian Stative: A Non-Finite Verb." *Ancient Near Eastern Studies* 53: 1-24. doi:10.2143/ANES.53.0.3154558.

Corbett, Greville G. 1991. *Gender*. Cambridge, MA: Cambridge UP.

Fischer, Wolfdietrich. 1997. "Classical Arabic." In *The Semitic Languages*, ed. by Robert Hetzron, 187-219. London-New York, NY: Routledge.

Fukui, Naoki. 2006. *Theoretical Comparative Syntax: Studies in Macroparameters*. London-New York, NY: Routledge.

Fuß, Eric. 2005. *The Rise of Agreement: A Formal Approach to the Syntax and Grammaticalization of Verbal Inflection*. Amsterdam-Philadelphia, PA: John Benjamins Publishing Company.

Getahun, Amare. 1996/1997 [1989]. *Zämänawi Yamariňňňa Säwasiw [A Modern Grammar of Amharic]*. Addis Abeba: Commercial Printing Press.

Greenberg, Joseph H. 1955. "Internal A-plural in Afroasiatic (Hamito-Semitic)." In *Afrikanistische Studien Diedrich Westerman Zum 80*, ed. by Jason H. Lucas, 198-204. Berlin: Berlin Akadamie Verlag.

Greenberg, Joseph H. 1991. "The Semitic 'Intensive' as Verbal Plurality: A Study of Grammaticalization." In *Semitic Studies in Honour of Wolf Leslau on the Occasion of His Eighty-Fifth Birthday (November 14th, 1991)*, ed. by Alan S. Kaye, vol. I., 577-587. Wiesbaden: Harrassowitz.

Gut, Ernest-August. 1997. "The Silte Group (East Gurage)." In *The Semitic Languages*, ed. by Robert Hetzron, 509-534. London-New York, NY: Routledge.

Hasselbach-Andee, Rebecca. 2007. "External Plural Markers in Semitic, A New Assessment." In *Studies in Semitic and Afroasiatic Linguistics Presented to Gene B. Gragg*, ed. by Cynthia L. Miller, 123-135. Chicago, IL, Ann Arbor, MI: the Oriental Institute, University of Chicago.

Hetzron, Robert. 1997. "The Outer South Ethiopic." In *The Semitic Languages*, ed. by Robert Hetzron, 535-549. London-New York, NY: Routledge.

Hudson, Grover. 1997. "Amharic and Argobba." In *The Semitic Languages*, ed. by Robert Hetzron, 457-485, London-New York, NY: Routledge.

Jastrow, Otto. 1997. "Neo-Aramaic." In *The Semitic Languages*, ed. by Robert Hetzron, 334-377. London-New York, NY: Routledge.

Kaufman, Alan S. 1997. "Aramaic." In *The Semitic Languages*, ed. by Robert Hetzron, 114-130. London-New York, NY: Routledge.
Kihm, Alain. 2008. "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*, ed. by Guglielmo Cinque and Richard S. Kayne, 459-512. Oxford: Oxford UP.
Kogan, Leonid E., and Andrey V. Korotayev. 1997. "Sayhadic (Epigraphic South Arabian)." In *The Semitic Languages*, ed. by Robert Hetzron, 220-241. London-New York, NY: Routledge.
Kouwenberg, N.J.C. 1997. "Gemination in the Akkadian Verb". *Studia Semitica Neerlandica* 33. Assen: Van Gorcum.
Kouwenberg, N.J.C. 2006. "The Proto-Semitic Gutturals in Old Assyrian." In *The Akkadian Language in its Semitic Context: Studies in the Akkadian of the Third and Second Millennium BC*, ed. by Guy Deutscher, and N.J.C. Kouwenberg, 150-176. Leiden: Netherlands Instituut voor Het Nabije Oosten.
Kramer, Ruth. 2014. "Gender in Amharic: A Morphosyntactic Approach to Natural and Grammatical Gender." *Language Sciences* 43: 102-115.
Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford UP.
Leslau, Wolf. 1995. *Reference Grammar of Amharic*. Wiesbaden: Harrassowitz.
Leslau, Wolf. 1997. *Ethiopic documents: Argobba Grammar and Dictionary*. Wiesbaden: Harrassowitz.
Lipiński, Edward. 1997. *Semitic Languages: Outline of Comparative Grammar*. Leuven: Uitgeverij Peeters.
Lowenstamm, Jean. 1991. "Vocalic Length and Syllable Structure in Semitic." In *Semitic Studies in Honour of Wolf Leslau on the Occasion of His Eighty-Fifth Birthday (November 14th, 1991)*, ed. by Alan S. Kaye, 949-965. Wiesbaden: Harrassowitz.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge UP.
Manzini, Rita M., Leonardo M. Savoia, and Tesfay Tewolde Yohannes. 2018. "Gender and Number Morphology in Tigrinya." In *International Conference on Eritrean Studies (20-22 July 2016, Proceedings)*, ed. by Zemenfes Tsige, Saleh Mahmud, Yonas Mesfun, Senai Woldeab, Rediet Kifle and Ghebreberhan Ogbazghi, 41-67. Asmara: Sabour Printing Services.
McCarthy, John. 1985. *Formal Problems in Semitic Phonology and Morphology*. London: Routledge.
Miyagawa, Shigeru. 2012. *Case, Argument Structure and Word Order*. New York, NY: Routledge.
Moscati, Sabatino, Anton Spitaler, Edward Ullendorff, and Wolfram Von Soden. 1964. *An Introduction to the Comparative Grammar of the Semitic Languages: Phonology and Morphology*. Wiesbaden: Harrassowitz.
Pardee, Dennis. 1997. "Ugaritic." In *The Semitic Languages*, ed. by Robert Hetzron, 131-144. London-New York, NY: Routledge.
Ratcliffe, Robert R. 1998. "Defining Morphological Isogloss: the 'Broken' Plural and Semitic subclassification." *Journal of Near Eastern Studies* 57 (2): 81-123.
Raz, Shlomo. 1983. *Tigre Grammar and Text*. Malibu, CA: Undena Publications.
Raz, Shlomo. 1997. "Tigre." In *The Semitic Languages*, ed. by Robert Hetzron, 446-456. London-New York, NY: Routledge.
Rubin, Aaron D. 2005. *Studies in Semitic Grammaticalization*: *Harvard Semitic Studies* 57. Winona Lake, IN: Eisenbrauns.
Sato, Yosuke. 2010. *Minimalist Interfaces*. Amsterdam-Philadelphia, PA: John Benjamins.
Siddiqi, Daniel. 2009. *Syntax Within the Word: Economy, Allomorphy, and Argument Selection in Distributed Morphology*. Amsterdam-Philadelphia, PA: John Benjamins.
Simpson, Andrew K. 2009. "The Origin and Development of Nonconcatenative Morpholgy." Ph.D. Dissertation. Berkeley, CA: University of California.
Steiner, Richard C. 1997. "Ancient Hebrew." In *The Semitic Languages*, ed. by Robert Hetzron, 145-173. London-New York, NY: Routledge.
Tesfay Tewolde Yohannes. 2002. *A Modern Grammar of Tigrinya.* Rome: Tipografia Ugo Detti.
Tesfay Tewolde Yohannes. 2003. "Plurals of Tigrinya and their Relationship." *Annali* 63: 83-98. Napoli: Università degli Studi di Napoli "l'Orientale".

Tesfay Tewolde Yohannes. 2009. "Relationship Among Verb Types and Internal Plurals in Eritrean Semitic." *Quaderni del Dipartimento di Linguistica* 19: 46-76.

Tesfay Tewolde Yohannes. 2016. *DPs, Phi-features and Tense in the Context of Abyssinian (Eritrean and Ethiopian) Semitic Languages*. Firenze: Firenze UP.

Tesfay Tewolde Yohannes. 2018. "Gender and Number Morphology in Tigrinya." In *International Conference on Eritrean Srudies (20-22 July 2016, Proceedings)*, ed. by Zemenfes Tsige, Saleh Mahmud, Yonas Mesfun, Senai woldeab, Rediet Kifle and Ghebreberhan Ogbazghi, 99-121. Asmara: Sabour Printing Services.

Wagner, Ewald. 1997. "Harari." In *The Semitic Languages*, ed. by Robert Hetzron, 486-508. London, New York, NY: Routledge.

Wälchli, Bernhard, and Francesca Di Garbi. 2019. "The Dynamics of Gender Complexity." In *Grammatical Gender and Linguistic Complexity*, ed. by Francesca Di Garbi, Bruno Olsson, and Bernhard Wälchli, 201-365. Berlin: Language Science Press.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

OPEN ACCESS

Citation: A. Wang (2021) *Wh*-phrases and *wh*-in-situ in Late Archaic Chinese. *Qulso* 7: pp. 155-179. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12008

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Wh-phrases and *wh*-in-situ in Late Archaic Chinese

Aiqing Wang
University of Liverpool (<aiqing.wang@liverpool.ac.uk>)

Abstract:

In this paper I explore *wh*-phrases and *wh*-in-situ in Late Archaic Chinese (LAC). Simplex and complex *wh*-phrases in LAC can be divided into eleven semantic categories. Since LAC is a *wh*-fronting language, *wh*-items undergo obligatory preposing, unless being subjects. Nonetheless, there are exceptions to the raising of non-subject *wh*-phrases, namely, obligatory and optional *wh*-in-situ. When *wh*-DPs function as the second complement of ditransitive verbs *nai*/*ruo*/*ru* 'to treat', or the second complement of the ditransitive verb *wei* 'to call', they must stay in their postverbal base position. In terms of optional *wh*-in-situ, there are two situations, i.e. *wh*-predicates and *wh*-complement of manner adverbials. When *wh*-DPs act as predicates indicating object/activity, person or reason, they normally do not move, yet they can move under special circumstances; when *wh*-VPs act as manner adjuncts, they can be base-generated pre- or postverbally, and the *wh*-elements in both positions undergo optional movement. Finally, I propose that *wh*-in-situ is correlated with ditransitive verbs per se.

Keywords: *Late Archaic Chinese, Syntax, Wh-in-situ, Wh-phrases*

1. Introduction

1.1 Historical context

Pre-Archaic (14thc BC-11thc BC) and Archaic Chinese (10thc BC-3rdc BC) exhibit robust syntactic features that are discrepant from those of Middle Chinese (2ndc BC-10thc AD) and modern Mandarin. For instance, the almost total lack of morphological marking of grammatical relationships is more salient during the (Pre-)Archaic period than at later stages. To be more specific, in Archaic Chinese: 1) lexical roots can be used freely as nouns, verbs or adjectives, and verbs can be used either transitively or intransitively without mediation of morphological marking; 2) nominalisation is employed to compensate the lack of finite em-

bedded clauses, and relative clauses on subject position and VP-internal positions are formed via separate strategies; and 3) there are different types of word order alternation, e.g. *wh*-movement, object focus fronting and pronoun preposing in the context of negation. Following Archaic Chinese, Middle Chinese appeared around the Han Dynasty (2ndc BC-2ndc AD) after the pre-Qin period, which was a crucial transitional period with multiple typological changes that were complete by the end of the Tang Dynasty, such as a rise of resultative compounds and an increase in embedded *wh*-questions. To be more specific, developments in Middle Chinese are: 1) loss of genitive case, as reflected by the mixed use of genitive and accusative pronouns and the fact that subjects of other types of embedded clause were not required to appear with genitive marking; 2) loss of subject/object relativisation asymmetry, caused by loss of a nominal layer in embedded clauses, which triggers the occurrence of a CP layer so that operators can move from either a subject or object position; 3) loss of (Pre-)Archaic Chinese movement transformations, triggered by loss of morphology for case and nominalisation; 4) disappearance of bare passives that were replaced by overtly marked passives, connected with loss of nominalising morphology; and 5) development of verb-resultative compounds, related to loss of causativising morphology. Changes took place in Middle Chinese are triggered by earlier morphosyntactic alternations and their subsequent loss (Xu 2006; Peyraube 2008; Aldridge 2013, 2015a).

In this paper I investigate the Classical Chinese par excellence, Late Archaic Chinese (LAC), which contains well-known texts such as *Analects*, *Mencius* and *Zhuangzi*.

1.2 Distinctive features of late archaic chinese

Texts in LAC display predominant SVO word order, with objects appearing in a postverbal position. However, there are contexts in which nominal and pronominal objects appear preverbally in the low TP-internal domain (Aldridge 2019), as in (1a-b).

(1) a. 吾 百姓 之 不 圖 (國語•越語下)
wu baixing$_i$ zhi bu [$_{VP}$ tu t$_i$]
I common.people ZHI not care.about
'I did not care about common people'

b. 吾 斯 之 未 能 信 (論語•公冶長)
Wu si$_i$ zhi wei neng [$_{VP}$ xin t$_i$]
I this ZHI not.yet can be.confident.in
'I am not confident in this yet'

Nonetheless, there is a robust disparity between LAC and modern Mandarin: differently from modern Mandarin which is a *wh*-in-situ language (Li 1992; Aoun and Li 1993, 2003; Tsai 1994, among many others), LAC requires VP-internal *wh*-phrases to raise from their base position to a preverbal position in the 'low IP area' (dubbed by Paul 2005) between TP and *v*P. Such clause-internal movement is driven by obligatory preverbal positioning of non-subject *wh*-elements of LAC, which is a *wh*-fronting language.

Examples (2a) and (2b) illustrate that both simplex *wh*-words and internally complex *wh*-phrases move to a preverbal position in the medial domain when acting as direct objects. In (2b), the nouns *battle*, and *alliance* are modified by a *wh*-word 何 *he* 'what', and they form a complex phrase preceding the *v*P.

(2) a. 然则 我 何 爲 乎? 何 不 爲 乎? (莊子•秋水)
Ranze wo he_i [$_{VP}$ wei t_i] hu? He_j bu [$_{VP}$ wei t_j] hu?
Then I what do Q what not do Q
'Then what do I do? What (do I) not do?'

b. 宋 何 役 之 不 會,
Song [he yi]$_i$ zhi bu [$_{VP}$ hui t_i],
Song what battle ZHI not enter
而 何 盟 之 不 同? (左傳•昭公二十五年)
Er [he meng]$_j$ zhi bu [$_{VP}$ tong t_j]?
Conj what alliance ZHI not join
'What battle does the State of Song not enter, and what alliance does (it) not join?'

In terms of subject *wh*-phrases, however, they remain in situ in LAC. When a *wh*-phrase occupies the subject position, it is in [Spec, TP], because Archaic Chinese has an A-position for the subject above *v*P.

1.3 Literature review

Notwithstanding examples (2-6) which exhibit preverbal objects, various observations support the view that LAC has always been an SVO language (Djamouri 2001; Djamouri and Paul 2009; Aldridge 2013; Djamouri, Paul and Whitman 2013), so object preposing is derived, and should not be assumed as the vestige of a basic OV word order, as proposed by Sun (1991), Feng (1996), Xu (2006), among others.

According to Aldridge (2012a, 2015b), non-*wh*-objects, including full NPs but excluding pronouns, always undergo syntactic focus movement into the low TP area, and obtain an interpretation of identificational focus. Meanwhile, topicalisation of non-pronominal DPs to the left periphery is common in LAC. With respect to the fronting of VP-internal *wh*-phrases, it is also limited to focus fronting, in that neither the base generation theory nor the cliticisation theory can explain the *wh*-fronting in LAC. Moreover, focalised *wh*-words are always located lower than modals and above negation (Aldridge 2006, 2007, 2010).

In terms of pronoun fronting in the context of negation, although the landing site of preposed pronouns intervenes between negation and *v*P, pronoun fronting to negation is not focus driven. A case-based approach has been put forward (Aldridge 2015b) to account for the motivation for pronoun fronting to negation in LAC: only pronouns in need of structural accusative case undergo fronting. As hypothesised by Aldridge (2015b), it is Neg that values accusative case on the fronted DPs, but the head of NegP selects a nominal complement *n*P where structural case is unavailable. Given the fact that *n* is a strong phase head and hence the unavailability of case in the domain of *n*, NP becomes impenetrable. As a consequence, DPs have to undergo object shift to [Spec, *n*P], so as to value accusative case from the head of NegP.

Paul (2002, 2005) suggests a parallelism between CP and the 'low IP area' in modern Mandarin, and she proposes a hierarchy 'CP > TopicP > even FocusP > IP > inner TopicP > even FocusP > *v*P'. In Modern Chinese, both TopicP and ModP are situated above the projection of 'even' focus (whether it is clause-internal or -external). Paul (2002, 2005) also argues that preposed objects occupy a specifier position of some functional category, instead of being adjoined to *v*P directly. Consequently, a preposed object does not occur in an adjoined position, but occupy the specifier node of a functional projection above the 'even' FocusP within the 'low IP area' (Paul 2005).

Nevertheless, Hsu's (2008) analysis on object preposing in modern Mandarin argues that the sentence-internal domain between TP and *v*P may not only license foci, but also topics. Given appropriate contexts, a preposed object can have either topic or focus status. However, instead of being topic or focus itself (Paul 2005), a preposed object requires two distinct projections for two interpretations. A topic and a focus occupy different functional projections, and they can co-occur in the sentence-internal domain, with the topic NP preceding the focus NP. Following Rizzi's (1997) "fine structure of the left periphery", Hsu (2008) posits that the functional projection for internal topics is located higher than that for foci.

Developing these lines of reasoning, Wang (2013, 2015) proposes an external topic position in the left periphery as well as an internal topic position and focus positions between TP and *v*P for the preposing of *wh*- and non-*wh*-objects in LAC, with the external and internal topic positions being structurally more prominent than the focus positions. All positions are located above negation. A fronted element targets the specifier node of some functional projection, followed by an optional fronting marker ZHI/SHI occupying the head of the corresponding category.

Peyraube and Wu (2005) discuss the origin and evolution of question words in Archaic Chinese and propound: 1) the specific forms of 曷 *he* 'when' tend to disappear towards the end of the LAC period; 2) as a general pronoun/adverb, 何 *he* 'what/which, how, why' exhibited more complex linguistic behaviour and a more abstract level of conceptualisation from Early Archaic Chinese to LAC; and 3) during the Archaic period, the system became more complex, as reflected by an increase of polysemy and synonyms. They have also proposed cognitive contours of conceptual categories of interrogative words: person > object > process/quality > space > time.

In this paper I analyse *wh*-phrases in LAC. In Section 2 I provide a descriptive account of eleven types of *wh*-words and their corresponding data drawn from corpora. In Section 3 I present examples involving *wh*-preposing as well as optional and obligatory *wh*-in-situ, and account for the *wh*-in-situ phenomenon in this *wh*-fronting language.

2. Types of wh-phrases

In LAC, *wh*-phrases display various forms, and can be divided into eleven semantic categories: 1) object/activity 'what/which', 2) person 'who', 3) manner 'how', 4) rhetorical 'how', 5) reason 'why', 6) instrument 'with what', 7) time 'when', 8) location 'where', 9) source 'from where', 10) direction 'to where' and 11) quantity 'how much/how many'. Each category contains a range of question words, and some question words are polysemous and appear in more than one group.

In the first category for object/activity, seven *wh*-words are attested: 何 *he*, 曷 *he*, 奚 *xi*, 胡 *hu*, 惡 *wu* and 孰 *shu*, as in (3a-f). Among these *wh*-words, 何 *he* is most frequently attested, and 曷 *he* is its variant. 奚 *xi* might be a variant of 何*he*, and both *he* and *xi* can function as a subject, object or attribute. 孰 *shu* and 奚 *xi* are different from their counterparts in this group, in that they convey the implication of choice, and the available options usually occur at the beginning of the sentence as antecedents (3f-g). That is to say, *shu* and *xi* mean 'which', whereas other *wh*-words in this group mean 'what' (Pulleyblank 1995).

(3) a. 然则 我 何 爲 乎? (莊子•秋水)
Ranze wo he_i [$_{VP}$ wei t_i] hu?
then I what do Q
'Then what do I do?'

b. 曷 謂 邪? (荀子•不苟)
He_i [$_{VP}$ wei t_i xie]
what call evil
'What (do we) call as evil?'

c. 孔子 奚 取 焉? (孟子•滕文公下; Peyraube 1997: 6)
Kongzi xi qu yan?
Confucius what approve.of in+him
'What did Confucius approve of in him?'

d. 胡 以 備 之? (管子•侈靡)
Hu_i [$_{VP}$ [$_{PP}$ yi t_i] bei zhi]?
what with prepare 3.Obj
'With what to prepare it?'

e. 女 將 惡 乎 比 予 哉? (莊子•人間世)
Ru jiang wu_i hu_j [$_{VP}$ bi yu [$_{PP}$ t_j t_i]] zai?
you Fut what to compare me Q
'To what will you compare me?'

f. 禮 與 食 孰 重? (孟子•告子下)
Li yu shi shu zhong?
etiquette Conj food which important
'Which is more important: etiquette or food?'

g. 其 一能 鳴, 其 一 不 能 鳴, 請 奚 殺? (莊子•山木)
Qi yi neng ming, qi yi bu neng ming, qing xi_i [sha t_i]?
between 1 can honk between 1 not can honk please which kill
'One (of the two geese) can honk, and the other cannot honk; which (goose should I) kill please?'

There is only one *wh*-word that falls into the group indicating person, namely, 誰 *shui* (4). 誰 *shui* has occurred since the period of Early Archaic Chinese (11th-6thc BC) and is still widely used as the main word for 'who' in modern Mandarin. 誰 *shui* in LAC may function as a subject, predicate, object or attribute.

(4) 誰 能 出 不 由 戶? (論語•雍也)
Shui neng chu bu you hu?
Who can go.out not through door
'Who can go out not through the door?'

The third group expressing manner 'how' contains both monosyllabic question words 何 *he*, 胡 *hu*, 安 *an*, 焉 *yan* and 奚 *xi* (5a-e) and disyllabic ones 何以 *heyi*, 奚以 *xiyi*, 惡乎 *wuhu*, 何如 *heru* and 何若 *heruo* (5f-j).

(5) a. 何 可 廢 也? (孟子•梁惠王上)
He ke fei ye?
How can stop Q
'How can (I) stop (it)?'

b. 胡 可 得 也? (新書)
Hu ke de ye?
How can realise Q
'How can (one) realise (it)?'

c. 余 安 能 知 之? (國語•周語中)
Yu an neng zhi zhi?
I how can know 3.Obj
'How can I know it?'

d. 焉 知 賢才 而 舉 之? (論語•子路)
Yan zhi xiancai er ju zhi?
How recognise talent Conj recommend 3.Obj
'How to recognise talents and recommend them?'

e. 奚 可 得 邪? (莊子•山木)
Xi ke de ye?
How can realise Q
'How can (one) realise (it)?'

f. 子 何以 知 之? (禮記•檀弓上)
Zi heyi zhi zhi?
You how know 3.Obj
'How did you know it?'

g. 奚以 益 之 而 治? (荀子•正名)
Xiyi yi zhi er zhi?
How benefit 3.Obj Conj govern
'How to benefit and govern it?'

h. 天下 惡乎 定? (孟子•梁惠王上)
Tianxia wuhu ding?
World how be.stable
'How (can) the world be stable?'

i. 德 何如,則 可以 王 矣? (孟子•梁惠王上)
De heru, ze keyi wang yi?
Virtue how then can be.king Q
'What kind of virtue can then make (one) king?'

j. 事 之 何若? (莊子•外物)
Shi zhi heruo?
make.progress 3.Obj how
'How is it going?'

All the above *wh*-words can be used in a rhetorical way and hence falling into the fourth category. The rhetorical interpretation of examples in (6a-j) is derived from context. Additionally, the question word 曷 *he* expressing object/activity (3b) can alternatively mean 'how' and be used in rhetorical sentences (6k).

(6) a. 文 王 何 可 當 也? (孟子•公孫丑上)
Wen wang he ke dang ye?
Wen King how can match.up Q
'How can King Wen be matched up?'

b. 又 胡 可 得 而 有 邪? (莊子•知北遊)
You hu ke de er you ye?
Then how can obtain Conj keep Q
'Then how can (one) obtain and keep (it)?'

c. 安 能 為 孝 乎? (禮記•祭義)
An neng wei xiao hu?
How can be filial.propriety Q
'How can (I be called a person of) filial propriety?'

d. 未 能 事 人, 焉 能 事 鬼? (論語•先進)
Wei neng shi ren, yan neng shi gui?
not.yet can serve people how can serve ghost
'(If you) cannot yet serve people, how can (you) serve ghosts?'

e. 臣 奚 能 言? (呂氏春秋•貴公)
Chen xi neng yan?
Subject how can say
'How can this subject (I) say?'

f. 吾 王 不 遊, 吾 何以 休? (孟子•梁惠王下)
Wu wang bu you, wu heyi xiu?
My king not travel I how rest
'How can I rest (if) my king does not travel?'

g. 奚以 知 其 然 也? (莊子•逍遙遊)
Xiyi zhi qi ran ye?
How know Gen correctness Q
'How (do they) know its correctness?'

h. 君子 去 仁, 惡乎 成 名? (論語•里仁)
Junzi qu ren, wuhu cheng ming?
Gentleman abandon benevolence how form reputation
'(If) gentlemen abandon benevolence, how (can they) form reputation?'

i. 何如 其 知 也? (論語•公冶長)
Heru qi zhi ye?
How Gen wisdom Q
'How (can one call it) his wisdom?'

j. 不 恥 不 若 人, 何若 人 有? (孟子•盡心上)
Bu chi bu ruo ren, heruo ren you?
Not be.ashamed.of not be.comparable.to people how people then
'(If one) is not ashamed of not being comparable to others, then how (can one catch up with) others?'

k. 天下 曷 敢 有 越 厥 志? (孟子•梁惠王下)
Tianxia he gan you yue jue zhi?
World how dare have disobey Gen will
'How dare anyone in the world disobey his will?'

The fifth group expressing reason is constituted of eleven mono- and disyllabic *wh*-words, viz. 何 *he*, 胡 *hu*, 盍(闔) *he*, 奚 *xi*, 焉 *yan*, 何以 *heyi*, 何故 *hegu*, 何為 *hewei*, 曷為 *hewei*, 奚以 *xiyi* and 奚為 *xiwei* (7). Among this group, 何 *he* and 胡 *hu* are the most common words for 'why', while 盍(闔) *he* is a newly emerged contracted form for 何 *he* and the negator 不 *bu*, indicating 'why not' (7c).

(7) a. 何 患 於 喪 乎 (論語•八佾)
He huan yu sang hu?
why upset by loss Q
'Why (are you) upset by the loss?'

b. 今 之 君子 胡 莫 行 之 也? (禮記•哀公問)
Jin zhi junzi hu mo xing zhi ye?
Today Gen gentleman why nobody conduct 3.Obj Q
'Why among today's gentlemen nobody conducts it?'

c. 子 盍 為 我 言 之? (孟子•公孫丑下)
Zi he wei wo yan zhi?
You why.not for me say 3.Obj
'Why don't you say it for me?'

d. 子 奚 不 為 政? (論語•為政)
Zi xi bu wei zheng?
You why not do politics
'Why do you not do politics?'

e. 子 為 政, 焉 用 殺? (論語•顏淵)
Zi wei zheng, yan yong sha?
You do politics why use killing
'When you do politics, why (do you) use killing?'

f. 何以 不 言 也? (公羊傳•隱公元年)
Heyi bu yan ye?
Why not say Q
'Why (did you) not say?'

g. 我 何故 不 得 福 也? (墨子•公孟)
Wo hegu bu de fu ye?
I why not receive blessing Q
'Why do I not receive blessing?'

h. 吾 何為 獨 不 然? (孟子•公孫丑下)
Wu hewei du bu ran?
I why alone not correct
'Why am I alone not correct?'

i. 曷為 三 遇 齊 王 而 不 言 事? (荀子•大略)
Hewei san yu Qi wang er bu yan shi?
Why thrice meet Qi king Conj not say thing
'Why (did you) meet the king of Qi thrice but not say anything?'

j. 奚以 之 九 萬 里 而 南 為? (莊子•逍遙遊)
Xiyi zhi jiu wan li er nan wei?
Why go 9 ten.thousand li (length unit) Conj south Q
'Why (do you) go (up for) ninety thousand li (a length unit) and then (fly) towards the south?'

k. 君 奚為 不 見 孟 軻 也? (孟子•梁惠王下)
Jun xiwei bu jian Meng Ke ye?
Your.Majesty why not meet Meng Ke Q
'Why did Your Majesty not meet Meng Ke?'

The next group expressing instrument involves two question words 何以 *heyi* and 奚以 *xiyi* (8a/b) that can mean 'how' (6f/g) or 'why' (7f/j) in other contexts.

(8) a. 將 何以 守 國? (國語•周語上)
Jiang heiyi shou guo?
Fut what.with guard state
'With what will (he) guard the state?'

b. 奚以 敬 民? (說苑•正諫)
Xiyi jing min?
what.with show.respect.to people
'What with (do we) show respect to people?'

The seventh category querying time information consists of one monosyllabic question word 曷 *he* 'when' and two disyllabic ones 何時 *heshi* and 奚時 *xishi* 'what time', as in (9a/b-c) respectively.

(9) a. 吾 子 其 曷 歸? (左傳•昭公元年)
Wu zi qi he gui?
My son part. When return
'When will my son return?'
b. 當 何時 作 之? (管子•度地)
Dang heshi zuo zhi?
Should what.time do 3.Obj
'What time should (we) do it?'
c. 而 人主 奚時 得 悟 乎? (韓非子•孤憤)
Er renzhu xishi de wu hu?
Conj monarch what.time can understand Q
'While what time can the monarch understand?'

As for the group expressing location, there are seven *wh*-words that fall into this group: 何 *he*, 安 *an*, 焉 *yan* and 惡 *wu*, as well as 何所 *hesuo*, 安所 *ansuo* and 惡乎 *wuhu* (10).

(10) a. 牛 何 之? (孟子•梁惠王上)
Niu he_i [$_{VP}$ zhi t_i]?
ox where go
'Where is the ox going?'
b. 安 在? (禮記•檀弓下)
An_i [$_{VP}$ zai t_i]?
where be.in
'Where is (him)?'
c. 將 焉 闢 之? (左傳•僖公九年)
Jiang yan bi zhi?
fut where avoid 3.Obj
'Where will (I) avoid it?'
d. 路 惡 在? (孟子•盡心)
Lu wu_i [$_{VP}$ zai t_i]?
road where be.in
'Where is the road?'
e. 子 何所 不 逞 欲 (左傳•昭公十四年)
Zi hesuo bu cheng yu?
you what.place not satisfy desire
'(In) what place do you not satisfy desires?'
f. 周 尚 安所 事 金 乎? (莊子•說劍)
Zhou shang ansuo shi jin hu?
Zhou then what.place use gold Q
'Then (in) what place does Zhou use the gold?'

g. 所謂, 惡乎 在? (莊子•知北遊)
Suowei dao, wuhu$_i$ [$_{VP}$ zai t$_i$]?
so-called Dao where be.in
'Where is the so-called Tao?'

The next category indicates source, and the question words are 焉 *yan* and 惡乎 *wuhu* (11).

(11) a. 而 君 焉 取 余? (左傳•莊公六年)
Er jun yan qu yu?
then Your.Majesty where obtain surplus
'Then (from) where does Your Majesty obtain the surplus?'
b. 惡乎 取 之? 取 之 曹 也。(公羊傳•僖公三十一年)
Wuhu$_i$ [$_{VP}$ qu zhi t$_i$]? Qu zhi Cao ye.
where take 3.Obj take 3.Obj Cao Decl
'(From) where to take it? Take it (from) Cao.'

In the group for direction, there are two interrogative words 焉 *yan* and 奚 *xi*, as in (12a) and (12b) respectively.

(12) a. 其 子 焉 往? (孟子•離婁上; Aldridge 2013: 246)
Qi zi yan wang?
3.Gen son where go
'Where will their sons go?'
b. 彼 且 奚 適 也? (莊子•逍遙遊)
Bi qie xi shi ye?
3.Obj then where go Q
'Where is it going?'

The last category consists of 幾 ji and 幾何 jihe, expressing 'how much/many'. The usage of 幾何 jihe has been preserved from Early Archaic Chinese (13a), whereas 幾 ji is a new expression that emerged during the LAC period (13b) (Peyraube and Wu 2000, 2005). As can be seen from (13a/13c) and (13b/13d) respectively, both jihe and ji can occur independently or be followed by a NP.

(13) a. 薛 之 地 小 大 幾何? (呂氏春秋•季冬紀)
Xue zhi di xiao da jihe?
Xue Gen land small big how.much
'How big is the land of Xue?'
b. 廢 者 幾? (禮記•曾子問)
Fei zhe ji
cease ZHE how.many
'How many (situations are there that are) ceased?'
c. 債 而 食 者 幾何 家? (管子•問第)
Zhai er si zhe jihe jia?
debt Conj feed ZHE how.many household
'How many households that (borrow) debt to feed?'

d. 子 來 幾 日 矣? (孟子•離婁上)
Zi lai ji ri yi?
you come how.many day Perf
'How many days have you been (here)?'

3. Wh-in-Situ and wh-preposing

Although LAC is a *wh*-fronting language, there are exceptions to the obligatory preposing of non-subject *wh*-items. In this section, I first give a descriptive account of non-subject *wh*-items that undergo obligatory fronting. Following that, I present non-subject *wh*-phrases that do not have to move and those that must stay in situ. Through comparing data in LAC and those in the following historical period and in modern Mandarin, I finally discuss the reasons of *wh*-in-situ in LAC.

3.1 Obligatory preposing of wh-phrases

As mentioned previously, LAC requires non-subject VP-internal *wh*-elements to raise from their base position to a preverbal position between TP and *v*P, so it is a *wh*-fronting language. Example (14a/b) show that when a *wh*-phrase occurs within *v*P, it has to move to a preverbal position across the verb, and this *wh*-element may refer to object/activity or person. As can be seen from (14b), an interrogative sentence is followed by its non-interrogative answer (in the form of a rhetorical question) with the identical verb, and the latter displays the canonical V-O order; since questions and their rhetorical counterparts are expected to share the same underlying structure, I state that the former interrogative sentence in (14b) involves *wh*-preposing and the surface *wh*-V order is generated via such *wh*-preposing.

(14) a. 然则 我 何 爲 乎? (莊子•秋水)
Ranze wo he_i [$_{VP}$ wei t_i] hu?
then I what do Q
'Then what do I do?'

b. 吾 誰 欺? 欺 天 乎? (論語•子罕)
Wu $shui_i$ [$_{VP}$ qi t_i]? qi tian hu?
I who deceive deceive Heaven Q
'Whom do I deceive? (Do I) deceive the Heaven?'

By contrast, when *wh*-elements function as adverbials and precede *v*P, they do not undergo (obligatory) movement, as in (15a-e) that involve question words indicating manner, rhetorical, reason, instrument and time.

(15) a. 余 安 能 知 之? (國語•周語中)
Yu an neng zhi zhi?
I how can know 3.Obj
'How can I know it?'

b. 吾 王 不 遊, 吾 何以 休? (孟子•梁惠王下)
Wu wang bu you, wu heyi xiu?
my king not travel I how rest
'How can I rest if my king does not travel?'

c. 子 奚 不 為 政? (論語•為政)
Zi xi bu wei zheng?
you why not do politics
'Why do you not do politics?'

d. 將 何 守 國? (國語•周語上)
Jiang heiyi shou guo?
fut what.with guard state
'With what will (he) guard the state?'

e. 當 何時 作 之 (管子•度地)
Dang heshi zuo zhi?
should what.time do 3.Obj
'What time should (we) do it?'

In terms of wh-phrases for location, source and direction, their (lack of) movement is correlated with their position, thus nature, in the sentence. When a *wh*-phrase precedes VP, it remains in its base position, as in (16). In examples involving in-situ *wh*-elements, verbs are transitive and take direct objects, while *wh*-phrases function as adverbials preceding the verb phrases.

(16) a. 子 何所 不 逞 欲? (左傳•昭公十四年)
Zi hesuo bu cheng yu?
you what.place not satisfy desire
'(In) what place do you not satisfy desires?'

b. 而 君 焉 取 余? (左傳•莊公六年)
Er jun yan qu yu?
then Your.Majesty where obtain surplus
'Then (from) where does Your Majesty obtain the surplus?'

c. 其 子 焉 往? (孟子•離婁上; Aldridge 2013: 246)
Qi zi yan wang?
3.Gen son where go
'Where will their sons go?'

Nevertheless, if these *wh*-phrases are base-generated postverbally within the *v*P, they must undergo preposing across the verb and land in a preverbal position, generating a *wh*-V order. Differently from *wh*-phrases in (16) that are adverbials, *wh*-phrases in (17) act as verbal complements. (17a/b/c) contains a pair of question and answer which are expected to demonstrate a parallel structure, which means the questions involve *wh*-movement from postverbal to preverbal positions. Assuming non-interrogative statement displays the canonical word order V-*wh*, the interrogative sentences in (17a-c) involve *wh*-preposing and thus a reverse *wh*-V order.

(17) a. 曰: '安 在?' 曰: '在 寢。' (禮記•檀弓下)
Yue: 'An$_i$ [$_{VP}$ zai t$_i$]?' Yue: 'Zai qin.'
say where be.in say be.in chamber
'(He) said: "Where is (him)?" (Someone) said: "(He) is in the chamber." '

b. 惡乎 取 之? 取 之 曹 也。 (公羊傳•僖公三十一年)
Wuhu$_i$ [$_{VP}$ qu zhi t$_i$]? Qu zhi Cao ye.
where take 3.Obj take 3.Obj Cao Decl
'(From) where to take it? Take it (from) Cao.'

c. 曰: '奚 之?' 曰: '將 之 衛。' (莊子•人間世)
Yue: 'Xi$_i$ [$_{VP}$ zhi t$_i$]?' Yue: 'Jiang zhi Wei.'
say where go say Fut go Wei
'(Confucius) said: "Where (are you) going?" (Yan Hui) said: "(I) will go to Wei." '

3.2 Obligatory wh-in-situ

There are two exceptions to the obligatory preposing of non-subject VP-internal *wh*-items: 1) the second complement of ditransitive verbs 奈/若/如 *nai*/*ruo*/*ru* 'to treat'; and 2) the second complement of the ditransitive verb 謂 *wei* 'to call'. *Wh*-DPs in these two situations must stay in situ.

First, in some double object constructions, if a *wh*-DP functions as the second complement, it must remain in situ. In LAC, there are three verbs that indicate 'to treat' and take two internal arguments, viz. *nai*/*ruo*/*ru*, and their second complement is always in situ if it is a *wh*-item (18a-c).

(18) a. 奈 吾 君 何? (國語•晋語二)
Nai [wu jun] he?
treat my lord what
'What does (this) do to my lord?'
b. 子 若 國 何? (左傳•僖公二十三年)
Zi ruo guo he?
you treat state what
'What do you do about the state?'
c. 將 如 君 何? (左傳•襄公二十三年)
Jiang ru jun he?
fut treat lord what
'What will (we) do to the lord?'

Second, when a *wh*-item acts as the second complement of a ditransitive verb *wei* 'to call', this *wh*-item must stay in situ. *Wei* may take a *wh*-phrase as its first or second complement, but there is an asymmetry between the two arguments of *wei*. When a *wh*-element acts as the first complement, it always fronts to a preverbal position (19a); by contrast, if a simplex or complex *wh*-phrase functions as the second complement, it normally remains in situ, as shown in (19b-c) and (19d) respectively, regardless of whether the first complement moves (19d) or not (19b-c).

(19) a. 何 謂 德義? (國語•晋語七)
He$_i$ [$_{VP}$ wei t$_i$] deyi
what call virtue.righteousness
'What (do we) call as virtue and righteousness?'
b. 國 謂 君 何? (左傳•僖公十五年)
Guo wei jun he?
state call lord what
'How does the state speak of the lord?' (Lit. 'What does the state call the lord?')
c. 吾 獨 谓 先 王 何 乎? (呂氏春秋•季秋紀)
Wu du wei xian wang he hu?
I alone call former lord what Q
'What do I alone call the former lord?'

d. 此 所 謂 何 聲 也? (韓非子•十過)
Ci$_i$ suo [$_{VP}$ wei t$_i$ [he sheng]] ye?
this SUO call what sound Q
'What sound (do we) call this?'

However, there is one and only one exception to the non-movement of the second complement: if and only if 1) the first argument fronts to a preverbal position, and 2) the second argument, i.e. the *wh*-phrase, is simplex, then the second argument raises to a position intervening between the fronted first complement and the ditransitive verb, as in (20a-b). The reason why examples in (19b-c) do not involve *wh*-movement is that in those two questions, the first object remains in situ, hence failing to meet the first condition. As for (19d), it contains a complex *wh*-phrase, hence failing to meet the second condition. Therefore, *wh*-items remain in situ in (19b-d).

(20) a. 是 何 謂 也? (左傳•昭公二十九年)
Shi$_i$ he$_j$ [$_{VP}$ wei t$_i$ t$_j$] ye?
this what call Q
'What (do we) understand these?'
b. 此 言 何 謂 也? (孟子•滕文公上)
[Ci yan]$_I$ he$_j$ [$_{VP}$ wei t$_i$ t$_j$] ye?
this sentence what call Q
'What (do we) call this sentence?'

It should be pointed out that the *wh*-in-situ phenomenon is not motivated by these ditransitive verbs per se.

First, I discuss the in-situ phenomenon concerning *nai*/*ruo*/*ru* 'to treat'. The first complement of *nai*/*ruo*/*ru* can raise out of the VP to a higher position, yet the movement of the first complement does not affect the non-movement of the second complement (*wh*-argument). In (21a), ZHI intervenes between the subject and verb to mark explicit subordination (Fuller 1999), and *Zhao Dun zhi fu guo* functions as the first argument of *nai* and raises to a preverbal position. The first argument in (21b) also moves out of the VP, but it also passes a modal verb and negation. As for (21c), it involves a hanging topic which is linked to a resumptive pronoun *zhi* as the first argument of the ditransitive verb, but again, this does not affect the second complement *he*. A *wh*-DP acting as the second argument in these three ditransitive constructions always remains in its base position, even if the first argument moves to a preverbal position (21a-b) or has a hanging topic (21c). Therefore, it is safe to say that the *wh*-in-situ is not caused by these three ditransitive verbs per se.

(21) a. 趙 盾 之 復 國 奈 何? (公羊傳•宣公六年)
[Zhao dun zhi fu guo]$_i$ [$_{VP}$ nai t$_i$ he]?
Zhao Dun ZHI recapture state treat what
'What (do we) do with Zhao Dun's recapture of the state?'
b. 知 其 不 可 奈 何 而 安 之 若 命 (莊子•人間世)
Zhi qi$_i$ bu ke [$_{VP}$ nai t$_i$ he] er an zhi ruo ming
know 3.Obj not can treat what Conj embrace 3.Obj as destiny
'(They) know there is nothing (they) can do about it, so (they) embrace it as the destiny.'
c. 君 使 臣, 臣 侍 君, 如 之 何? (論語•八佾)
[Jun shi chen, chen shi jun]$_i$, [$_{VP}$ ru zhi$_i$ he]?
lord employ official official serve lord treat 3.Obj what
'A lord employing officials and officials serving the lord, what about it?'

Second, *wh*-in-situ is not motivated by the ditransitive 謂 *wei* 'to call' per se. Providing the second argument of this ditransitive is a non-*wh*-constituent, it may raise to a preverbal position. For instance, in example (22) involving *wei*, the second object *bu shi shangxing* 'not neglect rewards and punishments' moves from its base position to a position preceding the ditransitive *wei*. Additionally, the first complement moves to the CP domain as a left-dislocated topic clause that is syntactically related to *wei* through linking to a gap, and this gap occupies the position of the first object. Therefore, I argue that the ditransitive verb *wei* itself does not prevent its second argument from raising, and the fact that some second arguments cannot raise is due to their interrogative nature (see (19b-d)).

(22) '惠 此 中 國, 以 綏 四方',
['Hui ci zhong guo, yi sui sifang']$_i$,
benefit this central state to appease four.direction
不 失 賞刑 之 謂 也 (左傳•僖公二十八年)
[bu shi shangxing]$_j$ zhi [$_{VP}$ wei t_i t_j] ye
Not neglect reward.punishment ZHI call Decl
'"Benefiting this central state to appease (vassal states in) all directions", (people) call it not to neglect rewards and punishments'

3.3 Optional wh-in-situ

There are two types of optional *wh*-in-situ, namely, *wh*-predicates and *wh*-complement of adjunct adverbials. *Wh*-predicates in LAC take the form of DPs and they indicate object/activity, person or reason. When *wh*-DPs function as predicates, they normally do not front, but they can raise under special circumstances. Moreover, *wh*-VPs functioning as manner adjuncts display flexible distribution: they can be base-generated either in a higher position above *v*P or in a lower postverbal position. The *wh*-elements in both preverbal and postverbal manner adverbials demonstrate optional movement. To be more specific, four types of constructions are attested: 1) V-*wh*-VP; 2) *wh*-V-VP; 3) VP-V-*wh*; and 4) VP-*wh*-V.

3.3.1 Wh-predicates

It is generally acknowledged that Archaic Chinese has no copula, so in equational sentences 'subject-nominal-*ye*', predicate nominals directly follow the subject, with a declarative particle 也 *ye* often occurring in a sentence-final position (Chang 2006). When *wh*-nominals function as predicates, they usually stay in situ, and they can question object/activity, person and reason.

First, a nominal predicate indicating object/activity is the simplex *wh*-word 何 *he* 'what' that follows the subject and stays in situ (23a-b). 何 *he* 'what' can combine with an NP and form a complex sentential predicate, as in (23c). As can be seen from (23a) and (23c), the declarative particle 也 *ye* may follow the sentential predicate *he* in a sentence-final position, but it is not obligatory (23b). Since *ye* typically accompanies a nominal predicate in LAC (Aldridge 2007), it serves as an (extra) piece of evidence justifying the predicate nature of *wh*-DPs following sentential subjects.

(23) a. 天下 之 害 何 也? (墨子•兼愛)
Tianxia zhi hai he ye?
world Gen calamity what Decl
'What is the world's calamity?'

b. 七 律 者 何? (國語•周語下)
[[Qi lv] zhe] he?
Seven rhythm ZHE what
'What are seven rhythms?'

c. 是 何 物 也? (左傳•昭公二十一年)
Shi [he wu] ye?
This what thing Decl
'What thing is this?'

Second, nominal predicates can ask for person information, indicated by the same simplex *wh*-word 何 *he* (meaning 'who' in this context) and another one 谁 *shui*. 何 *he* 'who' can appear independently (24a) or combine with an NP to form 何人 *he ren* 'what person' (24b), whereas 谁 *shui* can only be used alone (24c).

(24) a. 來 者 何? (穀梁傳•僖公四年)
[Lai zhe] he?
come ZHE who
'Who is the one that comes?'

b. 舜 何 人 也? (孟子•滕文公上)
Shun [he ren] ye?
Shun what person Decl
'What kind of person is Shun?'

c. 追 我 者 誰 也? (孟子•離婁; Aldridge 2007: 144)
[Zhui wo zhe] shui ye?
Pursue me ZHE who Decl
'Who is the one pursuing me?'

Third, simplex and complex *wh*-DPs indicating reason can serve as sentential predicates and stay in situ. As presented in Section 2, reason adverbials in LAC can be mono- or disyllabic, as in (25a) (=7d)) and (25b) (=(7g)) respectively. Reason adverbials are 'high' adjuncts in modern Mandarin (Tsai 2008); unsurprisingly, reason adverbials in LAC always occupy high positions in tree diagrams too. As opposed to these two 'high' adverbial constructions that are always base-generated above negation thus above *v*P, when simplex and complex *wh*-phrases function as predicates, they have the option to stay in their base position following the sentential subject. When the bare *wh*-word 何 *he* as well as complex *wh*-nominals 何故 *hegu* and 何義 *heyi* follow the sentential subjects as predicates, they remain in their original positions, as in (26). I treat simplex and complex *wh*-phrases following VP subjects (e.g. (26b)) as nominal predicates, following Peyraube and Wu (2000).

(25) a. 子 奚 不 為 政? (論語•為政)
Zi xi bu wei zheng?
you why not do politics
'Why do you not do politics?'

b. 我 何故 不 得 福 也? (墨子•公孟)
Wo hegu bu de fu ye?
I why not receive blessing Q
'Why do I not receive blessing?'

(26) a. 君 與 我 此 何 也? (國語•晉語一; Aldridge 2007: 144)
Jun yu wo ci he ye?
lord give me this what Decl
'Why is it that my lord gives me these things?'

b. 受 之 何 義 也? (孟子•章句下)
Shou zhi he yi ye?
accept 3.Obj what reason Decl
'(For) what reason (does one) accept it?'

c. 子 賀 我 何 故? (國語•晉語八)
Zi he wo he gu?
you congratulate me what reason
'(For) what reason do you congratulate me?'

It is important to point out that although the bare *wh*-word 何 *he* as well as complex *wh*-phrases 何故 *he gu* and 何義 *he yi* in (27) indicate reason, they cannot be treated as reason adverbials, as those in (25). First, according to my observation, only 何 *he* and complex phrases starting with 何 *he* can function as either predicative or adverbial phrases, yet the other *wh*-words such as 奚 *xi* (25a) can only act as adverbials. Second, as mentioned previously, *wh*-phrases functioning as adverbials of reason always appear as 'high' adverbials preceding *v*P (or more precisely, preceding negation), different from predicates that follow sentential subjects. Examples in (25) show that simplex and complex *wh*-adverbials are base-generated above negation. Moreover, examples in (27a-b) which are unmarked sentences with non-*wh*-PPs also help to show the base position of reason adverbials. Third, adverbials, but not predicates, can intervene between subject and *v*P. In (27c), if the adverbial 何 *he* was a predicate, *zi du fu zhi er xing* 'you alone carry them and walk' would be treated as a subject clause, but a predicate can never raise into a sentential subject and (27c) should not have been attested at all. Fourth, only under the analysis of *wh*-predicates can example (27d) be explained. Based on contextual information, 可 *ke* adopts the adjective meaning 'appropriate', so the *wh*-word 何 *he* has to act as a predicate rather than adverbial, otherwise this sentence would have no predicate. Moreover, movement of *wh*-phrases in LAC targets a node above negation (Aldridge 2006, 2007, 2010), which means LAC requires obligatory *wh*-preposing, and *wh*-phrases do not follow negation. If the *wh*-word 何 *he* in (27d) was a reason adverbial, it should have undergone *wh*-preposing to a position preceding the negator, but in the surface structure, the *wh*-word follows the negative. Therefore, the only justifiable account should be that 不可 *bu ke* functions as the sentential subject and the negation embedded within the subject cannot trigger the raising of the *wh*-predicate 何 *he*.

(27) a. 吾 以 故 知 古 從 之 同 也 (管子•白心)
Wu yi gu zhi gu cong zhi tong ye
I for reason understand ancient follow Gen similarity Decl
'For this reason I understand the ancient similarity of following'

b. 大王 若 以 此 不 信 (韓非子•難言)
Dawang ruo yi ci bu xin
your.Majesty if for this not keep.promise
'If Your Majesty does not keep promises because of this'

c. 今 子 何 獨 負 之 而 行? (韓非子•喻老)
Jin zi he du fu zhi er xing?
now you why alone carry 3.Obj Conj walk
'Now (for) what do you alone carry them and walk?'

d. 其 不 可 何 也 (韓非子•外儲說上)
Qi bu ke he ye?
M od not appropriate what Decl
'(For) what (is it) not appropriate?'

I state that *wh*-predicates involve optional *wh*-in-situ, because when serving as predicates, *wh*-phrases normally stay in situ, but they do undergo preposing under special circumstances. In (28), the simplex *wh*-DP *he* moves from a location following a sentential subject and lands in the left periphery as an external topic. In both examples, sentential subjects are nominalised by the morpheme 者 ZHE that can select a clausal complement projection and enable a TP to occupy an argument position as a sentential subject (Aldridge 2016).

(28) a. 何 哉 君 所 謂 逾 者? (孟子•梁惠王下)
He_i zai jun suo wei [yu zhe] t_i?
what Q Your.Majesty SUO call arrogate ZHE
'What is the arrogation that Your Majesty meant?'

b. 何 哉, 爾 所 謂 達 者? (論語•顏淵)
He_i zai, er suo wei [da zhe] t_i?
what Q you SUO call understand ZHE
'What is the understanding that you meant?'

3.3.2 Wh-complement of manner adverbials

In LAC, reason adverbials are 'high' adjuncts that always occupy high positions in trees, but non-reason manner adjunct adverbs can be base-generated either in a higher position above *v*P or be base-generated in a lower postverbal position. This observation coincides with the generalisation from Djamouri, Paul and Whitman (2013) that adjunct PPs in Classical Chinese may occur in a pre- or postverbal position.

Word order flexibility is a robust property of PPs in LAC, in that PPs can appear either before or after verbs (Peyraube 1996; Aldridge 2012b). Source PPs, for instance, display flexible distribution: in (29a), the preposition 自 *zi* 'from' and the nominal argument it selects precede the verb 反 *fan* 'to return', yet the source PP in (29b) occurs after the verb 出 *chu* 'to exit'.

(29) a. 世子 自 楚 反, 復 見 孟子 (孟子•滕文公上; Aldridge 2012b: 140)
Shizi zi Chu fan, fu jian Mengzi
heir from Chu return again see Mencius
'The heir returned from Chu and again visited Mencius'

b. 吾 聞 出 於 幽 谷 遷 于 喬 木 者 (*Ibidem*)
Wu wen chu yu you gu qian yu qiao mu zhe
I hear exit Loc dark valley move P tall tree Det
'I have heard of leaving a dark valley and heading to a tall tree'

When *wh*-phrases indicate manner, they also display flexible distribution and appear pre- or postverbally (as in (30-31)/(32-33) below), parallel to those expressing location, source and direction. Nonetheless, *wh*-phrases concerning manner are disparate from location, source and direction *wh*-phrases that adopt different features in different positions. To be more specific, as presented earlier in (16), when location, source and direction *wh*-phrases occur preverbally above *v*P, they function as adverbial adjuncts and do not move; by contrast, when location, source and direction *wh*-phrases occur postverbally within the *v*P, they must undergo preposing across the verb and land in a preverbal position, generating a *wh*-V order, because in this situation they act as VP-internal verbal complements (17). Manner *wh*-phrases, however, are always adverbial adjuncts, regardless of their locations.

Wh-phrases functioning as manner adjuncts can take two forms, viz. *wh*-DPs and *wh*-VPs. I hypothesise that when *wh*-VPs function as manner adverbials, they are always flexible in terms of *wh*-fronting (hence word order), regardless whether these manner *wh*-adverbials appear preverbally or postverbally. That is to say, there are four possible word orders, namely, V-*wh*-VP, VP-V-*wh*, *wh*-V-VP and VP-*wh*-V; the former two orders are canonical yet the latter two are generated via *wh*-raising.

Manner *wh*-phrases can take two forms, viz. *wh*-DPs and *wh*-VPs. Apart from simplex and complex *wh*-phrases 何 *he*, 胡 *hu*, 安 *an*, 焉 *yan*, 奚 *xi*, 何以 *heyi*, 奚以 *xiyi*, 惡乎 *wuhu* (see examples in (5)), manner adverbials in LAC can adopt another form, i.e. VP construction consisting of a ditransitive verb 奈/若/如 *nai*/*ruo*/*ru* 'to treat' and its *wh*-complement 何 *he* or 奚 *xi* 'what'. When a *wh*-VP functions as a manner adverbial, it can be base-generated preverbally. The *wh*-complement in the VP construction undergoes optional preposing, so that means both surface orders of [V-*wh*]-VP and [*wh*-V]-VP are attested. 何若 *he ruo* in (30a) is derived from *wh*-fronting out of the VP and the landing site of *he* is a preverbal position preceding the ditransitive verb *ruo*; the unmarked counterpart of (30a) is 若何 *ruo he* in (30b).

(30) a. 齊王 何 若 是 之 賢 也? (韓非子•外儲說下)
Qiwang he_i [$_{VP}$ ruo t_i] shi zhi xian ye?
Qi.emperor what treat this ZHI virtuous Q
'How can the Emperor of Qi (be) this virtuous?'

b. 若 何 許 之? (晏子春秋•內篇諫下)
[$_{VP}$ Ruo he] xu zhi?
treat what permit 3.Obj
'How (can you) permit him?'

When a manner adverbial is in the form of V-*wh* (30b), it seems that *wh* remaining in situ violates the requirement of obligatory *wh*-preposing in LAC. However, as discussed previously, *nai*/*ruo*/*ru* are ditransitives which require their second complement to remain in situ. I assume that 何 *he* 'what' in these VP constructions is the second complement, and the first complement is null. For instance, the unmarked, complete counterpart of 若何 *ruo he* (30b) is 若之何 *ruo zhi he* which contains both the first and the second complement and acts as a manner adverbial preceding *v*P (31).

(31) 若 之 何 不 弔? (左傳•襄公十四年)
[$_{VP}$ Ruo zhi he] bu diao?
Treat 3.Obj what not condole
'How (do we) not condole?'

I hypothesise that *wh*-in-situ in the surface structure of (30b) is caused by the fact that 何 *he* 'what' acts as the second argument of the ditransitive verb 若 *ruo*, thus prohibited from raising, even if being VP-internal. This presumption is supported by Peyraube and Wu (2000) that 若何 *ruo he* 'treat what' is derived from 若之何 *ruo zhi he* 'treat it what' from Early Archaic Chinese through the process of lexicalisation. Nevertheless, I notice that there are still instances involving *ruo zhi he* 'treat it what' in LAC period (as in (31)). With respect to examples that indeed involve *wh*-preposing like (30a), the explanation is that although adverbials in the V-*wh* order such as 若何 *ruo he* had developed into fixed expressions in the period of LAC, language users found the opposite order *wh*-V 何若 *he ruo* more acceptable based on their native speakers' intuition (Peyraube and Wu 2000). Therefore, *wh*-V constructions as in (30a) are still attested in LAC period.

Alternatively, *wh*-VP manner adverbials can be base-generated postverbally, and *wh*-complements within the VP constructions also have two options: they either front to a higher position across the verbs or stay in situ. Examples in (32a-c) involve ditransitive verbs 若 *ruo* and 如 *ru*, and the *wh*-complements they take can be 何 *he* or 奚 *xi*. I claim that examples in (32) involve fronting, because the canonical order between the verb and its argument has been inverted. Again, as discussed earlier, as the second complement of a ditransitive verb, *he* or *xi* is supposed to remain in situ. However, despite the fact that the V-*wh* order 如何 *ru he* and 若何 *ruo he* had developed into fixed expressions in LAC, language users found the reverse order *wh*-V more acceptable (Peyraube and Wu 2000). Therefore, *wh*-V constructions as in (32) still exist in corpora.

(32) a. 事 之 何 若? (莊子•外物)
Shi zhi he_i [$_{VP}$ ruo t_i]?
make.progress 3.Obj what treat
'How is it going?'

b. 當 皆 法 其 父母 奚 若? (墨子•法儀)
Dang jie fa qi fum xi_i [$_{VP}$ ruo t_i]?
if all emulate Gen parents what treat
'How is it if (people) all emulate their parents?'

c. 以 夫子 之 行為 奚 如? (莊子•天運)
Yi fuzi zhi xingwei xi_i [$_{VP}$ ru t_i]?
Think Confucius Gen behaviour what treat
'How (do you) think of Confucius' behaviour?'

There are other data illustrating that *wh*-complements of these ditransitive verbs can stay in their postverbal base position, and these V-O structures also function as manner adverbials (33).

(33) a. 而 為 之 若 何? (管子•侈靡)
Er wei zhi [$_{VP}$ ruo he]?
Conj conduct 3.Obj treat what
'Then how to conduct it?'

b. 佞人 之 事 君 如 何? (晏子春秋•內篇問上)
Ningrenzhi zhi shi jun [$_{VP}$ ru he]?
Sycophant ZHI wait.upon monarch treat what
'How do sycophants wait upon the monarch?'

To summarise, manner adverbials in the form of VPs exhibit flexible distribution and can be base-generated, pre- or postverbally. In both situations, *wh*-complements of ditransitive verbs within these manner adverbials are always free to either raise to a preverbal position and generate *wh*-V, or remain in situ thus the canonical order V-*wh*. In other words, *wh*-in situ concerning manner *wh*-adjuncts is optional.

3.4 Account of wh-in-situ

As mentioned previously, Archaic Chinese has no copula, which means *wh*-elements follow subjects directly without any linking elements in between. According to Aldridge (2006, 2007), *wh*-phrases in LAC move to a position between the subject and *v*P, rendering fronting of predicates pointless, so predicates do not raise in general, giving the appearance of *wh*-in-situ.

In terms of ditransitive verbs, the mismatch between their first argument that is subject to *wh*-fronting and their second argument that must stay in situ (cf. (34a) and (34b)) may be accounted by Stepanov's (2001, 2007) theory. If a thematic argument contains any uninterpretable feature (structural Case or *wh*-feature) in its label, it enters the structure by substitution, hence a structural argument; if not, the thematic argument enters the structure by adjunction, hence a structural adjunct. Only a structural argument is subject to movement, yet a structural adjunct without structural Case or *wh*-feature always enters the structure postcyclically. Since a structural adjunct cannot be Merged by the time the interrogative feature Q of the matrix complementiser is Merged with IP, as a consequence, an inherent Case marked DP cannot undergo raising and is inert. Note that the unmarked non-interrogative counterpart of (34a) is in the first clause of (34c).

(34) a. 何 謂 德義? (國語•晋語七)
He$_i$ [$_{VP}$ wei t$_i$] deyi
What call virtue.righteousness
'What (do we) call as virtue and righteousness?'

b. 國 謂 君 何? (左傳•僖公十五年)
Guo wei jun he?
state call lord what
'How does the state speak of the lord?' (Lit. 'What does the state call the lord?')

c. 夫 謂 之 辱 者, 非 此 之 謂 也 (呂氏春秋•正名)
Fu [wei zhi ru zhe]$_i$, fei ci$_j$ zhi [$_{VP}$ wei t$_i$ t$_j$] ye
Decl call 3.Obj humiliation ZHE FEI this ZHI call Decl
'That (we) call it humiliation, (we) do not call (it) as this'

In a ditransitive construction, the first complement is assigned accusative structural Case, yet the second argument receives dative inherent Case. A DP that is only marked inherent Case is inert (and transparent). For instance, a direct object in Albanian can move across an indirect object, but the indirect object cannot raise (McGinnis 1998; Stepanov 2007). Since both the first complement in (34a) and the second complement in (34b) display *wh*-feature, the only parameter that causes their disparity must be Case. The first thematic argument contains both *wh*-feature and structural Case, which makes it a structural argument, hence being subject to *wh*-fronting. However, the second thematic argument only has *wh*-feature, yet the inherent Case it receives brings the derivational property of inertness. Consequently, extraction out of an inherently Case marked DP is impossible.

With respect to the situations of obligatory *wh*-in-situ, they can be accounted for by the related verbs per se.

First, the second complement of ditransitive verbs *nai*/*ruo*/*ru* 'to treat' must stay in situ in LAC. I argue that these constructions were employed as fixed expressions in that historical period only, which explains the fact that such ditransitive use does not exist in modern Mandarin anymore.

Second, the second complement of the ditransitive verb *wei* 'to call' must stay in situ, and I argue that this is also due to the verb per se and is correlated with the Intervention Effect. The Intervention Effect as in Beck (1996) and Beck and Kim (1997) denotes the fact that a barrier may not intervene between a question existential operator (Q-operator) and a function variable bound by that Q-operator. The Intervention Effect in Mandarin is triggered by focus constructions, and it can be circumvented by a repair strategy raising an in-situ *wh*-item to a position preceding the focus-induced barrier (Kim 2002, 2006). I use (35a) that contains a ditransitive verb *gei* 'to give' to illustrate the Intervention Effect in Mandarin, and (35a) displays the canonical order. In (35b), if the first complement *Lisi* is focalised and fronts to a preverbal position, the Q-binding of the second complement, the *wh*-phrase *na-ben shu* 'which book', will be blocked by this focus barrier; as a consequence, (35b) becomes infelicitous, because the *wh*-element is preceded by a focalised object and subject to the Intervention Effect. (35c) and (35d) demonstrate that the *wh*-DP must move to a higher position across the focus-induced barrier in order to be bound by its operator, and the fronted *wh*-DP can move to the CP domain (35c) or stay in the low IP area (35d), as long as its landing site is more prominent than the focus-induced barrier.

(35) a. Zhangsan gei Lisi na-ben shu?
Zhangsan give Lisi which-CL book
'Which book does Zhangsan give Lisi?'

b. *Zhangsan lian $Lisi_i$ dou [VP gei t_i [na-ben shu]]?
Zhangsan even Lisi also give which-CL book

c. [Na-ben shu]$_j$ Zhangsan lian Lisi dou [$_{VP}$ gei t_i t_j]?
which-CL book Zhangsan even Lisi also give

d. Zhangsan [na-ben shu]$_j$ lian $Lisi_i$ dou [$_{VP}$ gei t_i t_j]?
Zhangsan which-CL book even Lisi also give
'Which book does Zhangsan give even Lisi?'

I propose that due to the language per se, the ditransitive verb *wei* 'to call' is even stronger than the Intervention Effect, in a sense that *wei* can suppress the Intervention Effect and allow its second complement (*wh*-DPs) to stay in situ, without having to undergo raising across focus-induced barriers. Consequently, owing to the robustness of *wei*, it is natural for its second complement to remain in situ under normal circumstances without the Intervention Effect. Due to insufficiency of data concerning *wei* in LAC corpora, I resort to its modern counterpart of *jiao* (36a). As shown in (35), the second complement of ditransitive verbs like *gei* is affected by preceding focus constructions, so the *wh*-DP has to move across the blocking element in order to circumvent the Intervention Effect. *Wh*-DPs introduced by *jiao*, however, are immune from the Intervention Effect. Therefore, *shenme* 'what' in (36b) can remain in its base position, despite the fact that there is a focused subject intervening between *shenme* and its Q-operator. Since the modern counterpart of *wei* is strong enough to circumvent the Intervention Effect, it is unsurprising to see that *wei* in LAC can enable its *wh*-complement to stay in situ.

(36) a. Zhangsan [$_{VP}$ jiao Lisi shenme]?
Zhangsan call Lisi what
'What does Zhangsan call Lisi?'
b. Lian Zhangsan dou [$_{VP}$ jiao Lisi shenme]?
even Zhangsan also call Lisi what
'What does even Zhangsan call Lisi?'

4. Conclusion

In this paper, I analyse *wh*-phrases in LAC, focusing on situations when they stay in situ. Analogous to modern Mandarin, LAC is an SVO language; yet differently from its modern counterpart, LAC requires non-subject VP-internal *wh*-elements to undergo obligatory clause-internal preverbal positioning, driven by the fact that LAC is a *wh*-fronting language. Nevertheless, there are data involving obligatory and optional *wh*-in-situ attested in this historical period (5thc BC-3rdc BC).

Before discussing *wh*-preposing and *wh*-in-situ, I give a descriptive account of eleven categories of simplex and complex *wh*-phrases and their corresponding data drawn from corpora. Question words in LAC can be divided into eleven semantic groups, indicating object/activity, person, manner, rhetorical, reason, instrument, time, location, source, direction and quantity. Each category is constituted of a range of different question words, and there are overlaps among groups, because it is common for (question) words to be polysemous in LAC, which might be related to lack of inflectional morphology.

LAC entails obligatory preposing of non-subject VP-internal *wh*-items, but there are two exceptions to such movement. First, if a *wh*-DP acts as the second complement of ditransitive verbs 奈/若/如 *nai*/*ruo*/*ru* 'to treat', it normally stays in situ, unless this *wh*-DP is simplex and the first argument fronts to a preverbal position. Second, if a *wh*-DP acts as the second complement of the ditransitive verb 謂 *wei* 'to call', it must stay in situ following the verb. I then point out that the obligatory *wh*-in-situ under these two circumstances is not motivated by these ditransitive verbs per se.

Apart from obligatory *wh*-in-situ, there are two situations when *wh*-elements do not have to undergo obligatory preposing, yet they do not have to stay in situ either. The first type of optional *wh*-in-situ is *wh*-predicates. When *wh*-nominals function as predicates, they can question object/activity, person or reason. A *wh*-predicate usually stays in situ, unless this predicate is 何 *he* 'what' and the raising takes place in the format of [$_{TP}$ Subj-SUO-V$_1$-[V$_2$-ZHE]]-*wh*-*zai*; under this particular circumstance, the surface structure becomes *wh*-*zai*-[$_{TP}$ Subj-SUO-V$_1$-[V$_2$-ZHE]]. The second type of optional *wh*-in-situ is the *wh*-complement inside *wh*-VPs that act as manner adverbials. When *wh*-VPs function as manner adjuncts, they show flexible distribution and can be base-generated either above *v*P or inside *v*P. Whether the VP adverbials being preverbal or postverbal, the *wh*-elements inside these VPs always undergo optional movement. That is to say, a manner adverbial may display a canonical order V-*wh* or a surface order *wh*-V (generated via *wh*-fronting); in either order, the manner adverbial can precede or follow the VP.

In the following section, I explain *wh*-fronting in LAC by referring to the theory of Stepanov (2001, 2007). I also suggest that obligatory *wh*-in-situ concerning *wei* is correlated with this ditransitive verb per se. Since the modern counterpart of *wei* enables its *wh*-complement to be immune from the Intervention Effect (Beck 1996; Beck and Kim 1997), I assume that *wei* can also suppress the Intervention Effect, and due to its robustness, it is not surprising for its complement to remain in situ without the Intervention Effect.

References

Aldridge, Edith. 2006. "VP-internal Quantification in Old Chinese." In *Chinese Linguistics in Budapest*, vol. I, *Chinese Linguistics in Europe*, ed. by Redouane Djamouri, and Rint Sybesma, 1-15. Paris: Ecole des Hautes Etudes en Sciences Sociales, Centre de Recherches Linguistiques sur l'Asie Orientale.

Aldridge, Edith. 2009. "*Wh*-indefinites and their Relation to *Wh*-in-Situ." In *Proceedings from the Panels of the Forty-third Annual Meeting of the Chicago Linguistic Society: Volume 2*, vol. XLIII, ed. by Malcolm Elliott, James Kirby, Osamu Sawada, Eleni Staraki, and Suwon Yoon, 139-153. Chicago, IL: Chicago Linguistic Society.

Aldridge, Edith. 2010. "Clause-internal Wh-movement in Archaic Chinese." *Journal of East Asian Linguistics* 19 (1): 1-36. doi: 10.1007/s10831-009-9054-z.

Aldridge, Edith. 2012a. "Focus and Archaic Chinese Word Order." In *The Proceedings of the 22nd North American Conference of Chinese Linguistics (NACCL-22) and the 18th Annual Meeting of the International Association of Chinese Linguistics (IACL-18)* (Cambridge, MA, 20-22 May 2010), vol. II, ed. by Lauren E. Clemens, and Chi-Ming L. Liu, 84-101. Columbus, OH: The Ohio State University.

Aldridge, Edith. 2012b. "PPs and Applicatives in Late Archaic Chinese." *Studies in Chinese Linguistics* 33 (3): 139-164.

Aldridge, Edith. 2013. "Survey of Chinese Historical Syntax: Pre-archaic and Archaic Chinese." *Language and Linguistics Compass* 7 (1): 58-77.

Aldridge, Edith. 2015a. "Wh-questions, Premodern." In *Encyclopedia of Chinese Language and Linguistics*, ed. by Rint Sybesma. Leiden: Brill.

Aldridge, Edith. 2015b. "Pronominal Object Shift in Archaic Chinese." In *Syntax over Time: Lexical, Morphological and Information-structural Interactions*, ed. by Theresa Biberauer, and George Walkden, 350-370. Oxford: Oxford UP.

Aldridge, Edith. 2016. "Relativization and DP Structure in Late Archaic Chinese." In *New Horizons in the Study of Chinese: Dialectology, Grammar, and Philology-Studies in Honor of Professor Anne Yue*, ed. by Pang-hsin Ting, Samuel Hung-nin Cheung, Sze-Wing Tang, and Andy Chi-on Chin, 429-446. Hong Kong: T.T. Ng Chinese Language Research Centre, Chinese University of Hong Kong.

Aldridge, Edith. 2019. "Subject/non-subject Movement Asymmetries in Late Archaic Chinese." *Glossa: A Journal of General Linguistics* 4 (1): 115. doi: 10.5334/gjgl.743.

Aoun, Joseph, and Y.H. Audrey Li. 1993. "Wh-elements in Situ: Syntax or LF?" *Linguistic Inquiry* 24 (2): 199-238.

Aoun, Joseph, and Y.H. Audrey Li. 2003. *Essays on the Representational and Derivational Nature of Grammar: The Diversity of Wh-Constructions.* London-Cambridge, MA: MIT Press.

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: De Gruyter.

Beck, Sigird. 1996. "Quantified Structures as Barriers for LF Movement." *Natural Language Semantics* 4 (1): 1-56.

Beck, Sigrid, and Shin-Sook Kim. 1997. "On Wh- and Operator Scope in Korean." *Journal of East Asian Linguistics* 6 (4): 339-384.

Chang, Jung-hsing. 2006. "The Chinese Copula *Shi* and Its Origin: a Cognitive-based Approach." *Taiwan Journal of Linguistics* 4 (1): 131-156.

Djamouri, Redouane. 2001. "Markers of Predication in Shang Bone Inscriptions." In *Sinitic Grammar: Synchronic and Diachronic Perspectives*, ed. by Hilary Chappell, 143-171. Oxford: Oxford UP.

Djamouri, Redouane, and Waltraud Paul. 2009. "Verb-to-preposition Reanalysis in Chinese." In *Historical Syntax and Linguistic Theory*, ed. by Paola Crisma, and Giuseppe Longobardi, 194-211. Oxford: Oxford UP.

Djamouri, Redouane, Waltraud Paul, and John Whitman. 2013. "Syntactic Change in Chinese and the Argument-adjunct Asymmetry." In *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*, vol. II, ed. by Guangshun Cao, Hilary Chappell, Redouane Djamouri, and Thekla Wiebusch, 577-594. Taipei: Academia Sinica.

Feng, Shengli. 1996. "Prosodically Constrained Syntactic Changes in Early Archaic Chinese." *Journal of East Asian Linguistics* 5 (4): 323-371.

Fuller, Michael A. 1999. *An Introduction to Literary Chinese*. Cambridge, MA: Harvard UP.

Hsu, Yu-Yin. 2008. "The Sentence-Internal Topic and Focus in Chinese." In *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20), Volume 2* (Columbus, OH, 25-27 April 2008), ed. by Marjorie K. M. Chan, and Hana Kang, 635-652. Columbus, OH: The Ohio State University.

Huang, C.T. James, Y.H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge-New York, NY: Cambridge UP.

Kim, Shin-Sook. 2002. "Intervention Effects Are Focus Effects." In *Japanese/Korean Linguistics 10*, ed. by Noriko Akatsuka, and Susan Strauss, 615-628. Stanford, CA: Center for the Study of Language and Information.

Kim, Shin-Sook. 2006. *Intervention Effects in Questions*. Ph.D. Dissertation, University of Frankfurt.

Li, Y.H. Audrey. 1992. "Indefinite Wh in Mandarin Chinese." *Journal of East Asia Linguistics* 1 (2): 125-156.

McGinnis, Martha. 1998. *Locality in A-movement*. Ph.D. Dissertation. Cambridge, MA: MIT Press.

Paul, Waltraud. 2002. "Sentence-internal Topics in Mandarin Chinese: The Case of Object Preposing." *Language and Linguistics* 3 (4): 695-714. Taipei: Academia Sinica.

Paul, Waltraud. 2005. "Low IP Area and Left Periphery in Mandarin Chinese." *Recherches Linguistiques de Vincennes* 33: 111-134.

Peyraube, Alain. 2008. "Ancient Chinese." In *The Ancient Languages of Asia and the Americas*, ed. by Roger D. Woodard, 136-162. Cambridge: Cambridge UP.

Peyraube, Alain. 1996. "Recent Issues in Chinese Historical Syntax." In *New Horizons in Chinese Linguistics*, ed. C.T.J. Huang and Y-H. A. Li, 161-214. Dordrecht: Kluwer.

Peyraube, Alain, and Fuxiang Wu. 2000. "Shanggu hanyu yiwen daici de fazhan yu yanbian." [*Development and Evolution of The Interrogative Pronouns in Chinese*]. *Zhongguo yuwen* 4: 311-326.

Peyraube, Alain, and Fuxiang Wu. 2005. "Origin and Evolution of Question-words in Archaic Chinese: A Cognitive Approach." *Cahiers de Linguistique-Asie Orientale* 34 (1): 3-24.

Pulleyblank, Edwin G. 1995. *Outline of Classical Chinese Grammar*. Vancouver: University of British Columbia Press.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery." In *Elements of Grammar*, ed. by Liliane M.V. Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.

Sagart, Laurent. 1999. *The Roots of Old Chinese*. Amsterdam-Philadelphia, PA: John Benjamins.

Stepanov, Arthur. 2001. "Late Adjunction and Minimalist Phrase Structure." *Syntax* 4 (2): 94-125.

Stepanov, Arthur. 2007. "The End of CED? Minimalism and Extraction Domains." *Syntax* 10 (1): 80-126.

Sun, Chaofen. 1991. "The Adposition Yi and Word Order in Classical Chinese." *Journal of Chinese Linguistics* 19 (2): 202-219.

Tsai, Wei-Tien D. 1994. *On Economizing The Theory of A-bar Dependencies*. Ph.D. dissertation. Cambridge, MA: MIT Press.

Tsai, Wei-Tien D. 2008. "Left Periphery and How-why Alternations." *Journal of East Asian Linguistics* 17 (2): 83-115.

Wang, Aiqing. 2013. "Object Preposing in Late Archaic Chinese." *York Papers in Linguistics* 2 (13): 70-98.

Wang, Aiqing. 2015. "The Intervention Effect of Negation on Wh-adverbials in Late Archaic Chinese." *The Second Asian and European Linguistic Conference Proceedings*, ed. by Alison Biggs, Man Li, Aiqing Wang, and Cong Zhang, 160-172.

Wu, Fuxiang. 1980. *Gudai hanyu* [Classical Chinese]. Lanzhou: Gansu Renmu Chubanshe.

Xu, Dan. 2006. *Typological Change in Chinese Syntax*. Oxford: Oxford UP.

Fonologia

Phonology

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: L. Bafile (2021) Struttura X-barra nei segmenti: la rappresentazione della lenizione. *Qulso* 7: pp. 183-203. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12058

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Struttura X-barra nei segmenti: la rappresentazione della lenizione

Laura Bafile
Università di Ferrara (<laura.bafile@unife.it>)

Abstract:

The article concerns some developments of Element Theory that extend X-bar theory to the realm of phonology in order to represent the internal structure of segments. While maintaining some fundamental assumptions of Element Theory concerning the nature of melodic primes, these models propose a radically different conception of the segment and of larger phonological domains compared to autosegmentalist approaches. The article discusses some issues emerging from the adoption of X-bar theory and notation with regard to the representation of lenition phenomena, and suggests that the complex hierarchical structure that according to these theories corresponds to a segment poses problems of descriptive adequacy and raises a general question about learnability.

Keywords: *Element Theory, Lenition, Recursion, Segmental phonology*

1. Introduzione

La concezione secondo cui i primitivi segmentali sono *elementi*, cioè unità monovalenti e dotate di un'autonoma interpretazione fonetica, ha determinato progressi significativi nella capacità della teoria fonologica di spiegare aspetti statici (proprietà dei domini fonologici) e dinamici (fenomeni di cambiamento) dell'organizzazione fonologica. I vantaggi di questo approccio, rispetto ai modelli che fanno riferimento ai tratti binari della fonologia generativa standard, sono particolarmente evidenti nell'ambito dei fenomeni di lenizione e indebolimento vocalico (Kaye, Lowenstamm e Vergnaud 1985; Harris 1994; Harris e Lindsey 1995, 2000; Backley 2011).

Il termine *lenizione* è comunemente usato in fonologia per indicare un insieme di processi segmentali, complessivamente caratterizzati come "indebolimento", che interessano consonanti diverse, in contesti diversi e con esiti vari. Tale complessità empirica e descrittiva costituisce un'evidente difficoltà per un

approccio teorico alla questione e in effetti, per quanto a un livello di analisi descrittivo sia una nozione comunemente utilizzata, la teoria della lenizione costituisce un tema controverso (si veda Honeybone 2008 per una rassegna storica dei concetti di forza consonantica e lenizione).

In particolare, per la Teoria degli Elementi (*Element Theory,* ET) nella sua formulazione classica (Harris 1994; Harris e Lindsey 1995, 2000), la rappresentazione dei fenomeni di lenizione e di indebolimento vocalico ha avuto un ruolo cruciale nella costruzione stessa della teoria, oltre a costituire uno dei principali campi di analisi. I modelli costruiti a partire dalla ET classica hanno mantenuto la concezione fondamentale dei primitivi fonologici ma hanno modificato significativamente, radicalmente in certi casi, aspetti rilevanti della teoria, dapprima modificando l'inventario degli elementi e poi arrivando a un completo cambiamento della concezione stessa di segmento.

In questo articolo prenderemo in considerazione un particolare sviluppo della ET che propone per il livello segmentale una struttura essenzialmente identica a quella sintattica, in cui il contenuto segmentale è organizzato in costituenti di tipo X-barra. Di questi due modelli, denominati *GP 2.0* e *Precedence-Free Phonology* (PFPh), esamineremo alcuni aspetti pertinenti alla rappresentazione della lenizione.

Il paragrafo 2 contiene una sintesi degli aspetti fonologici della lenizione e nel paragrafo 3 sono presentati i punti essenziali della Government Phonology e della Teoria degli Elementi. Il paragrafo 4 presenta i punti essenziali delle revisioni della ET classica che sono state proposte negli anni e si sofferma su due modelli in particolare, GP 2.0 (paragrafo 4.1) e Precedence-Free Phonology (paragrafo 4.2). Seguono alcune riflessioni conclusive.

2. Aspetti fonologici della lenizione

Il termine *lenizione* è usato con una varietà di significati. Nella tradizione dialettologica italiana indica il processo sincronico che colpisce le consonanti occlusive sorde, e in certa misura anche le altre ostruenti sorde, in posizione intervocalica, nei dialetti italiani centro-meridionali e nelle corrispondenti varietà regionali di italiano. Questo fenomeno di lenizione può essere descritto come una parziale sonorizzazione (cfr. Loporcaro 1988: 105-112 per una sintesi; Nocchi e Schmid 2008 soprattutto sugli aspetti fonetici). Più in generale, nella letteratura fonologica, soprattutto in lingua inglese, *lenizione* indica un insieme di processi che colpiscono diverse consonanti o classi di consonanti, in contesti diversi, con esiti diversi (su vari aspetti fonologici di lenizione e rafforzamento cfr. Carvalho, Ségéral e Scheer 2008). Nella descrizione di Hyman (1975: 165) (che attribuisce la definizione a Theo Vennemann) l'indebolimento consonantico (Hyman non usa qui il termine *lenition*) è un processo che trasforma un segmento in un altro che corrisponde a uno stadio più avanzato di un percorso di riduzione a zero. Una trafila di indebolimento è ad esempio la seguente (Hyman 1975: 164):

(1) tappu > tapu > tabu > taβu > tawu > tau > to:

In (1) si osservano scempiamento, sonorizzazione, spirantizzazione, sonorantizzazione e cancellazione.

Le consonanti più tipicamente coinvolte dalla lenizione sono le ostruenti sorde e tra queste le occlusive. Tuttavia, mentre in alcune lingue è un'intera serie a subire l'indebolimento (come nel caso della già menzionata lenizione italiana centro-meridionale) in altre lingue la lenizione interessa specifici segmenti distinti in base al luogo di articolazione. Questo tipo di asimmetrie ha portato Foley (1970) a proporre una gerarchia di forza consonantica rela-

tiva al luogo di articolazione, ordinando labiali > dentali > velari; ma questa ipotesi è stata messa in discussione sulla base di dati che mostrano che, laddove il luogo di articolazione è discriminante, la relativa gerarchia varia da lingua a lingua (cfr. Hyman 1975: 166-167; Kirchner 1988: 7).

Per quanto riguarda il contesto, la lenizione colpisce consonanti che si trovano in posizione intervocalica o in posizione di coda sillabica. Anche da questo punto di vista si osserva variazione interlinguistica. I sistemi interessati da lenizione possono presentarne gli effetti solo tra vocali (e questo è il caso della maggior parte dei fenomeni osservabili sincronicamente nei sistemi italoromanzi) o solo in posizione di coda; alcune lingue presentano esiti leniti sia tra vocali sia in coda. Ad esempio, la spirantizzazione del toscano e la semisonorizzazione dei dialetti centro-meridionali, per cui ad esempio /dato/ → da[θ]o (fiorentino), da[ʒ̬]o (romanesco), è limitata alla posizione intervocalica.[1] Al contrario, in molte varietà di inglese, /t/ subisce lenizione tanto tra vocali quanto in coda. Ad esempio, nell'irlandese meridionale e nell'inglese del Merseyside (contea di Liverpool) /t/ viene spirantizzata in entrambe le posizioni, come in [ˈlɛsə] 'letter', [ˈgɛs] 'get', mentre la varietà di inglese di New York City e di altre zone degli Stati Uniti e della Gran Bretagna presenta *flapping* in posizione intervocalica e perdita di rilascio in posizione di coda o finale assoluta, per cui pi[ɾ]y 'pity', fi[ɾ] us 'fit us', ma fi[tˀ] me 'fit me' (Harris 1994: 121).

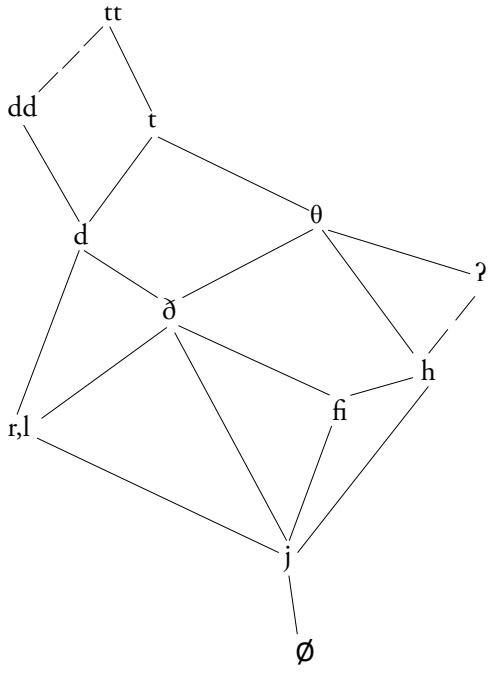
Riguardo all'*output* della lenizione, gli esiti possibili per una stessa consonante sono plurimi, non solo nel senso evidenziato dall'esempio in (1), che illustra stadi progressivi dell'indebolimento di un'occlusiva. Una stessa occlusiva può alternare con esiti vari per effetto di cambiamenti del contenuto fonetico di tipo diverso, come perdita di occlusione, perdita di rilascio, debuccalizzazione, sonorizzazione, o una combinazione di essi. Ad esempio:

(2)		
	spirantizzazione	t → θ
	perdita di rilascio	t → tˀ
	sonorizzazione (parziale o totale)	t → ʒ̬, d
	perdita di rilascio + debuccalizzazione	t → ʔ
	spirantizzazione + debuccalizzazione	t → h
	sonorizzazione + spirantizzazione	t → ð

La varietà degli esiti fa ben vedere come una rappresentazione monodimensionale della lenizione sia una inevitabile semplificazione. Ad esempio, secondo Hock (1991: 83), un modello adeguato richiede una rappresentazione multidimensionale complessa (le linee tratteggiate rappresentano sviluppi non documentati):

[1] Ai fini della lenizione centro-meridionale e della "gorgia toscana", la posizione intervocalica include anche i contesti V_lV, V_rV, V_wV, V_ jV.

(3)



Tale complessità, e la problematicità del concetto di *forza segmentale* a cui il concetto di indebolimento fa ovviamente riferimento, sono affrontati in un modello classico della lenizione quale quello di Lass (1984: 177 e segg.) che contempera due approcci in genere alternativi, individuando nell'indebolimento consonantico due possibili traiettorie, quella della *apertura* (cfr. Lass e Anderson 1975) e quella della *sonorità* (nel senso dell'inglese *sonority*). I due percorsi coincidono in buona parte, cogliendo entrambi il passaggio occlusiva > fricativa > approssimante, ma, secondo Lass, la disgiunzione è necessaria per rappresentare da un lato il completo dileguo (massima apertura laringale e sopralaringale, ma abbattimento della sonorità), e dall'altra il passaggio sorda > sonora (incremento di sonorità, ma non di apertura). La proposta di Lass non riesce però a risolvere i problemi che il concetto stesso di forza presenta in generale e in particolare per la rappresentazione della lenizione (cfr. Lavoie 2001). Il concetto di apertura coglie direttamente i cambiamenti riguardanti il modo di articolazione, ma dà una spiegazione meno convincente della debuccalizzazione. Come osserva Harris (2009), tale approccio dovrebbe riuscire a dimostrare che una restrizione nel cavo orale sia più costrittiva di una restrizione glottidale, il che è forse possibile per il passaggio [s] → [h], ma improbabile per [t] → [ʔ].

Una fenomenologia così complessa dal punto di vista descrittivo pone delle ovvie difficoltà a un approccio teorico esplicativo. Al di là dei problemi di rappresentazione, superabili con l'adozione di un formalismo sufficientemente efficace, la questione teorica essenziale emerge negli approcci ispirati a un principio epistemologico di non arbitrarietà, per cui la spiegazione dei fenomeni fonologici deve esplicitare la relazione esistente tra cambiamento e contesto.

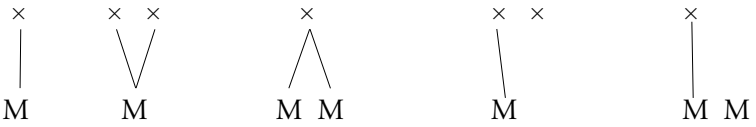
Una concezione di questo tipo è quella che caratterizza la *Government Phonology,* di cui la ET fa parte (Kaye, Lowenstamm e Vergnaud 1990; Harris 1994, 1997). Nel prossimo paragrafo

prenderemo in esame la teoria della lenizione sviluppata nella Government Phonology nella sua versione classica, ed esamineremo dalla prospettiva della spiegazione dei fenomeni di lenizione alcuni sviluppi successivi proposti all'interno dello stesso quadro teorico.

3. Government Phonology e la rappresentazione della lenizione

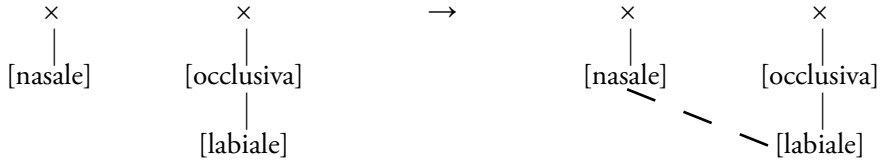
Government Phonology (GP) è una teoria fonologica sviluppata a partire dalla concezione autosegmentale della rappresentazione fonologica (Goldsmith 1976), secondo cui le diverse proprietà delle espressioni fonologiche sono organizzate su livelli autonomi anche se collegati tra loro. Il cardine di tale organizzazione è l'asse delle posizioni temporali '×', a cui si associano da un lato il livello melodico, cioè quello del contenuto segmentale, dall'altro il livello dell'organizzazione prosodica, a partire dai costituenti sillabici.[2] Nel modello autosegmentale, l'associazione di una posizione temporale con un certo contenuto melodico, costituito da primitivi subsegmentali (tratti o elementi), corrisponde alla nozione tradizionale di segmento. Rispetto al segmento dei modelli unilineari, la rappresentazione autosegmentale prevede una corrispondenza non necessariamente univoca tra contenuto melodico (M) e posizione ×, che include anche la possibilità di una posizione temporale vuota o di un contenuto melodico non associato:

(4)



Il modello autosegmentale è in sé un presupposto per rappresentazioni non arbitrarie dei fenomeni fonologici, in quanto esplicita la relazione esistente tra contesto e fenomeno rappresentato. Ad esempio, dato un passaggio come [np] → [mp], la rappresentazione autosegmentale esplicita il carattere assimilatorio e la causa contestuale del fenomeno (vedi (5)), che resterebbero invece inespressi in una rappresentazione fonologica lineare. Quest'ultima infatti, con gli stessi strumenti formali, può rappresentare senza differenze regole fonologicamente motivate e regole completamente arbitrarie, come ad esempio *n* → *m* /__ *k*.

(5)



I problemi di arbitrarietà insiti nella fonologia generativa standard (Chomsky e Halle 1968) erano noti e da più parti evidenziati (cfr. ad esempio Hooper 1976; Kiparsky 1982; Kaye 1989; Harris 1994).

In GP, la non arbitrarietà delle rappresentazioni fonologiche assume il valore di un principio essenziale. Le possibilità offerte dal modello autosegmentale vengono ulteriormente formalmente ristrette da fondamenti epistemologici che si traducono negli elementi costitutivi

[2] Nei sistemi tonali, anche le unità tonali sono collocate su un asse autonomo, indipendente da quello melodico e da quello prosodico.

della rappresentazione stessa: un principio di località, che lega indissolubilmente i fenomeni segmentali al loro contesto; un principio di conservazione, secondo cui la struttura sillabica lessicale è mantenuta inalterata nel corso della derivazione; il principio in base al quale le espressioni fonologiche sono interpretabili dalla fonetica in tutti i livelli della derivazione e quello per cui i processi fonologici si applicano ogni qualvolta si presentino le condizioni contestuali utili (per un'efficace sintesi sui fondamenti epistemologici di GP si veda Kaye 2005, Pöchtrager e Kaye 2013). In questo impianto teorico, che limita drasticamente quanto può avvenire nella derivazione fonologica, ha un ruolo fondamentale una sottoteoria relativa ai segmenti del tutto innovativa rispetto alla fonologia generativa standard, cioè la ET, che integra GP fin dalle sue prime formulazioni, (Kaye, Lowenstamm e Vergnaud 1985, 1990). Nella ET le unità subsegmentali sono tratti, ovvero *elementi,* monovalenti e dotati di un proprio corrispettivo fonetico, cioè autonomamente pronunciabili. In questa visione i segmenti sono dei composti di elementi che come tali possono subire solo processi di scomposizione e composizione, con perdita o redistribuzione di elementi già presenti nella rappresentazione lessicale (cfr. Kaye, Lowenstamm e Vergnaud 1985).

Nella versione della ET che definirò "classica" (Harris 1994; Harris e Lindsey 1995, 2000) gli elementi sono 10: gli elementi "di risonanza", che portano il contenuto relativo al timbro nelle vocali e al luogo di articolazione nelle consonanti, gli elementi di modo di articolazione e gli elementi che portano il contenuto relativo all'attività della laringe:

(6)

risonanza	**modo**	**laringe**
A I U R@	h, ʔ, N	H L

Ciascun elemento ha una diretta interpretazione fonetica che emerge quando l'elemento è l'unico componente di una espressione segmentale. In combinazione con altri elementi, le proprietà acustiche dell'elemento possono determinare effetti fonetici diversi a seconda che l'espressione segmentale sia associata a un nucleo o a un attacco. La tabella in (7) illustra sinteticamente il sistema previsto nella versione classica della ET; non saranno esaminate qui in ulteriore dettaglio le caratteristiche del sistema illustrato in (7), alcune delle quali problematiche e controverse, per cui si rimanda alla trattazione di Backley (2011).

(7)

	interpretazione autonoma	**nelle V**	**nelle C**
A	[a]	struttura formantica di [a]	luogo faringale
I	[i]	struttura formantica di [i]	luogo palatale
U	[u]	struttura formantica di [u]	luogo labiale
R	[ɾ]		luogo coronale
@	[ə]	struttura formantica di [ə]	luogo velare
ʔ	[ʔ]		occlusione
h	[h]		rilascio turbolento
N	nasale	nasalità	nasalità
L		tono basso	sonorità
H		tono alto	assenza di sonorità

Dato l'inventario di elementi in (7), ad esempio, una consonante occlusiva bilabiale corrisponde all'espressione in (8).

(8) [U ʔ h] [p]

Per quanto riguarda le caratteristiche relative allo stato della laringe, la descrizione tradizionale che individua un contrasto sonoro/sordo non rappresenta adeguatamente la complessità empirica. In una parte delle lingue, la proprietà pertinente e fonologicamente attiva (ad esempio nelle assimilazioni) è la sordità, mentre il membro "sonoro" dell'opposizione è in realtà neutro, cioè privo di proprietà. In altre lingue, al contrario, la proprietà attiva è la sonorità, e il membro "sordo" della coppia è l'elemento neutro. Nelle restanti lingue, infine, entrambe le proprietà sono fonologicamente pertinenti e attive. Nella ET, questa variazione relativa allo stato della laringe dipende dagli elementi **H** (sordità) e **L** (sonorità) (Harris 1994; Honeybone 2008; Backley 2011). Quella in (8) è la rappresentazione di un'occlusiva sorda di una lingua come l'italiano, in cui le ostruenti sorde sono neutre, mentre le ostruenti sonore contengono **L**.

Per completare questa descrizione molto sintetica dei punti salienti della teoria, osserviamo che il concetto di primitivi segmentali come unità privative e singolarmente pronunciabili è connesso fin dalle prime teorizzazioni con la nozione di *dipendenza*, secondo cui gli elementi all'interno di un segmento possono avere pesi diversi, contribuendo in modo ineguale alla sua composizione (cfr. Anderson e Jones 1974; Schane 1984; Kaye, Lowenstamm e Vergnaud 1985; Harris e Lindsey 1995). Nella ET la dipendenza è formalmente espressa attraverso la nozione di *testa.* In un'espressione segmentale un elemento indicato come testa apporta il contenuto prevalente; ad esempio, l'espressione [**A I**], in cui **A** è testa (segnalata dalla sottolineatura), corrisponde a una vocale bassa [æ], mentre [**A I**] corrisponde a una vocale medio-alta [e].

Come già osservato, per la ET i fenomeni di indebolimento segmentale, diffusamente osservabili nelle lingue del mondo, rivestono un'importanza cruciale. Essi infatti costituiscono una base empirica di prioritaria importanza per la concezione secondo cui i segmenti sono dei composti di unità monovalenti dotati di un'autonoma interpretazione fonetica. In questa prospettiva, la lenizione può essere rappresentata in modo molto semplice attraverso un processo di erosione del contenuto segmentale e perdita degli elementi; per la stessa ragione, esiti di lenizione diversi e alternativi possono essere rappresentati come perdita di elementi diversi. Ad esempio, data la composizione di un'occlusiva bilabiale come rappresentata in (8), gli esiti di lenizione possono corrispondere alle espressioni segmentali in (9 a-e):

(9)		[U ʔ h]		
	a.	[U ʔ]	pˀ	perdita di rilascio
	b.	[ʔ]	ʔ	glottalizzazione
	c.	[U h]	ɸ	spirantizzazione
	d.	[h]	h	spirantizzazione + debuccalizzazione
	e.	[U]	w	vocalizzazione

Un altro aspetto di rilevanza teorica è la definizione dei contesti di lenizione. Nell'ambito della GP, la questione è stata specificamente trattata all'interno della *Lateral Phonology* (Scheer 2004). Più specificamente, la teorizzazione nota come *Coda Mirror* (Ségéral e Scheer 2001, 2008; Scheer 2004) offre una soluzione al problema di circolarità insito nella descrizione dei contesti di lenizione e rafforzamento, secondo cui "l'indebolimento si applica nelle posizioni deboli (coda sillabica); il rafforzamento si applica nelle posizioni forti (iniziale (assoluta) e postconsonantica)". *Coda Mirror* mette in una connessione formale tali processi con il contesto in cui si applicano. Un ulteriore problema è costituito dal fatto che la lenizione si applica in contesti diversi, cioè in posizione di coda e in posizione intervocalica. Su questo punto la GP

classica non dà una soluzione (al di là della generica osservazione che la consonante si trova in un contesto di massima apertura) e la Lateral Phonology può dare una soluzione di tipo stipulativo (cfr. Bafile 2019 per una discussione). Il rapporto tra fenomeni di indebolimento vocalico e consonantico e la struttura prosodica al di sopra della sillaba è trattato in Harris (1994, 1997) all'interno del modello *Licensing Inheritance*, secondo cui la capacità di una posizione × di sostenere contenuto segmentale complesso dipende dalla forza prosodica della posizione stessa: una posizione è più forte se corrisponde a un nucleo accentato o dipende da un nucleo accentato anziché da un nucleo atono.

Nel resto di questo articolo ci occuperemo solo degli aspetti della lenizione connessi con il contenuto segmentale delle consonanti coinvolte.

4. Sviluppi della Teoria degli Elementi

La ET classica (Harris 1994; Harris e Lindsey 1995, 2000) è stata oggetto nel corso degli anni di revisioni in alcuni casi molto profonde, che ne hanno modificato vari aspetti, a partire dall'inventario degli elementi. Le motivazioni alla base di queste revisioni sono fondamentalmente di due tipi. Da un lato vi è l'obiettivo di un'adeguata rappresentazione delle categorie fonologiche pertinenti nella composizione dei segmenti; in questa prospettiva, l'esatta forma fonetica dei suoni ha un'importanza limitata, dato che il contenuto fonologico dei segmenti è rivelato dal loro comportamento all'interno degli enunciati oltre che dalle caratteristiche acustiche essenziali (cfr. Kaye 2005). Il secondo argomento è invece interno alla teoria, ed è quello che fa riferimento al pericolo della "*overgeneration*". Il termine, riferito a modelli di sistemi fonologici, indica la capacità di generare insiemi di segmenti molto più grandi degli inventari effettivamente osservabili nelle lingue naturali. Nell'ambito della ET, tale questione ha ricevuto fin dall'inizio (cfr. ad esempio Backley 1993) notevole attenzione, e non è un caso che quasi tutte le proposte di revisione dell'inventario menzionino tra le motivazioni più importanti il rischio di *overgeneration* (cfr. Bafile 2020 per una discussione).

Le proposte di modifica dell'inventario dei segmenti hanno preso in considerazione le diverse classi di elementi. Per quanto riguarda il luogo di articolazione, già Harris e Lindsey (1995) riconoscevano la problematicità di **R** e **@**: la caratteristica dell'elemento neutro è quella di essere una "non proprietà", cioè di essere privo di proprietà fonologicamente pertinenti. Per quanto riguarda la coronalità, da un lato le caratteristiche acustiche corrispondenti a **R** sono sfuggenti, dall'altro il comportamento fonologico delle consonanti coronali è notoriamente particolare in molte lingue, in quanto queste consonanti mostrano diffusamente un comportamento passivo, nel senso che tendono a essere suoni inattivi o trasparenti in processi fonologici, come ad esempio l'assimilazione o l'armonia vocalica (cfr. Paradis e Prunet 1991; Backley 1993; Harris e Lindsey 1995: 68-69).

Backley (2011) rappresenta la versione definita "standard", che ha accolto alcune delle proposte di revisione presentate fin dagli esordi della ET. In merito agli elementi di luogo, gli elementi **@** e **R** sono soppressi. Labialità e velarità convergono e vengono rappresentate entrambe con l'elemento **U** (**U** testa = labiale, **U** non testa = velare) (Backley 2011: 79). La coronalità viene ad essere espressa dall'elemento di palatalità **I** oppure dall'elemento di faringalità **A** (Backley 2011: 77-79). Secondo l'autore, l'osservazione empirica evidenzia da una parte che le coronali sono caratterizzate in alcune lingue da **I** e in altre da **A**, e dall'altra parte che, in una stessa lingua, alcune coronali possono comportarsi come segmenti contenenti **I** e altre come segmenti contenenti **A**. L'eliminazione di **R** persegue anche un obiettivo più generale, cioè quello di una rappresentazione completamente unificata del contenuto melodico di consonanti

e vocali nel rispetto del principio che le categorie melodiche non sono di per sé consonantiche o vocaliche, e che ciò che fa di un composto segmentale una consonante o una vocale è il suo status sillabico, cioè l'associazione, rispettivamente, a un attacco (o una coda) o a un nucleo.

Quest'ultima considerazione vale anche per gli elementi di modo, il secondo aspetto su cui si sono incentrate le proposte di revisione. La ET standard prevede la fusione di **h**, l'elemento esclusivamente consonantico contenuto nelle occlusive e nelle fricative, con **H**. Nell'analisi di Backley e Nasukawa (2009) e Backley (2011), **H** caratterizza il tono alto nelle vocali, è contenuto nelle consonanti occlusive e nelle fricative in cui corrisponde al "rumore" e, in lingue come l'inglese, determina la presenza di aspirazione nelle occlusive e l'assenza totale di sonorità nelle ostruenti più in generale. L'elemento di occlusione è invece mantenuto in Backley (2011), in quanto giustificato anche in associazione a nuclei sillabici dall'esistenza di vocali glottalizzate, rappresentate come composti contenenti **ʔ** (torneremo su questo punto in 4.2).

Il terzo cambiamento riguarda la rappresentazione della nasalità, che nella ET standard converge con la sonorità nell'elemento **L**. Dal punto di vista acustico, il "mormorio" caratteristico dei suoni nasali e il segnale glottidale della sonorità consistono entrambi in energia periodica a bassa frequenza. In una prospettiva tipologica, si osserva che molte lingue sono caratterizzate da restrizioni fonotattiche o processi che evidenziano uno stretto legame tra le occlusive sonore e le occlusive nasali, chiaramente distinte, come insieme, dalle occlusive sorde (cfr. Nasukawa 2005).

In sintesi, la ET standard contiene i sei elementi **A I U H L ʔ**.

Altre teorizzazioni hanno proposto revisioni più radicali, che non riguardano solo la codifica delle proprietà melodiche, ma anche la natura stessa dei segmenti. Nei prossimi paragrafi prenderemo in esame in particolare due modelli che propongono una rappresentazione di tipo sintattico della struttura dei segmenti.

4.1 GP 2.0

La teoria fonologica nota come GP 2.0 (Pöchtrager 2006, 2020; Pöchrager e Kaye 2013; Živanović e Pöchtrager 2010) si pone dichiaratamente in continuità con la Government Phonology classica di cui mantiene alcuni caratteri fondamentali. La concezione del segmento, tuttavia, è profondamente rivista.

Rispetto all' inventario degli elementi, GP 2.0 propone una riduzione rispetto al sistema a sei unità della ET standard (Backley 2011). Oltre alla motivazione di carattere generale già menzionata, cioè il pericolo di *overgeneration,* la proposta di eliminare alcuni elementi muove da considerazioni di tipo empirico.

Pöchtrager (2006) riprende un problema sollevato da Jensen (1994) riguardo alla natura dell'elemento **ʔ**. Tale elemento presenta infatti caratteristiche anomale rispetto agli altri primitivi melodici. Una prima considerazione è di carattere interno alla teoria, e si riferisce al fatto che la distribuzione di **ʔ** è ristretta alle sole posizioni consonantiche, e in tal modo è impedita una generalizzazione desiderabile già menzionata, secondo cui tutti gli elementi possono associarsi sia a posizioni consonantiche sia a posizioni vocaliche (ricordiamo però che secondo Backley 2011 **ʔ** è presente nelle vocali laringalizzate). La seconda considerazione di Jensen (1994) si riferisce a fenomeni di alternanza come quelli osservabili nel fula (o pulaar), una lingua dell'Africa occidentale (cfr. ad esempio Anderson 1976). Nel fula, all'interno di parola le consonanti non occlusive, se soggette a geminazione per ragioni morfologiche, vengono sostituite da un corrispondente occlusivo, ad esempio nell'alternanza *lew-ru* / *lebbi* < **lew-wi* 'mese/mesi' (Pöchtrager 2006: 40). Inoltre, il fula presenta in posizione iniziale la cosiddetta *gradazione consonantica,* un'alternanza tra consonanti occlusive e non occlusive, ad esempio *waa-ndu* / *baa-di* 'scimmia/

scimmie' (Anderson 1976), la cui distribuzione dipende interamente da fattori morfologici, a differenza di quanto si osserva in lingue imparentate in cui l'alternanza è condizionata fonologicamente in quanto dipendente da condizioni contestuali. Il problema è quindi spiegare la comparsa di **ʔ** all'interno di una delle due consonanti che alternano senza che tale elemento sia disponibile nel contesto.

Jensen (1994) ipotizza che la comparsa di **ʔ** sia collegata a una precisa configurazione strutturale, cioè a una struttura coda-attacco corrispondente a una geminata, tanto in posizione interna di parola, in cui la geminata è realizzata come tale, quanto in posizione iniziale, in cui la geminata è virtuale (consistente in una struttura parzialmente vuota). L'occlusività non è quindi una caratteristica melodica del segmento, ma una caratteristica strutturale.

Pöchtrager (2006) estende questa argomentazione anche all'elemento **H**, proponendo di cancellarlo dall'inventario dei primitivi melodici. Questa idea prende spunto dall'analisi delle alternanze di lunghezza vocalica dell'inglese di New York City. Oltre al contrasto di lunghezza (che si accompagna a una differenza di qualità vocalica), come ad esempio in *beat* / *bit* [iː] / [ɪ], questa varietà di inglese presenta sistematiche differenze di lunghezza dipendenti dalla sonorità della ostruente successiva (Pöchtrager 2006: 17 e sgg). Ad esempio:

(10)	*bit*	*beat*	*bid*	*bead*
	[ɪ]	[iː]	[ɪː]	[iːː]

In (10) si osserva quindi che se la consonante finale non contiene l'elemento **H** (che determina la sordità delle consonanti in lingue come l'inglese), la vocale precedente si allunga. Apparentemente, quindi, una caratteristica melodica, cioè sordità/non sordità, si intreccia stranamente con una caratteristica strutturale, cioè la lunghezza, che nella rappresentazione autosegmentale è collocata su un livello autonomo.[3] Pöchtrager (2006) ne trae la conclusione che non esiste un'unità melodica **H** e che anche questa proprietà dei segmenti, oltre all'occlusività, è codificata nella struttura. Nella teorizzazione di Pöchtrager, la struttura necessaria a codificare tutte le informazioni pertinenti è generata da un'applicazione reiterata dell'operazione Merge. Questa rappresentazione non prevede la multilinearità propria della rappresentazione autosegmentale; l'informazione melodica (elementi) e sillabica (attacco, nucleo) è direttamente annotata sulle posizioni × che formano la struttura ai vari livelli di proiezione.

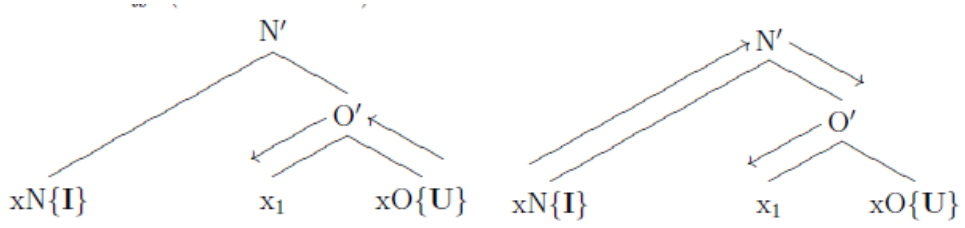
In (11) è rappresentata la struttura della coppia di fricative *f* / *v* dell'inglese (da Pöchtrager 2006: 66; 'O' sta per *onset*):

(11) a. [v] b. [f]

[3] Considerazioni analoghe sono applicate da Pöchtrager (2006) all'analisi della lunghezza vocalica in estone. Non ci soffermiamo qui su questo punto.

Quella indicata dalla freccia in (11b) è la relazione di m-comando esercitata dalla posizione di attacco sulla posizione × precedente. L' attacco di (11b) è "più lungo" perché, attraverso l'm-comando, occupa due posizioni ×. La differenza tra (11a) e (11b) è stabilita lessicalmente. La posizione $\times_1$ in (11a) resta disponibile e può essere occupata, tramite m- comando, da un nucleo precedente. Questo è ciò che accade in questa varietà di inglese in parole come *leave*[iː] e *whiff* [ɪ] (da Pöchtrager 2006: 69; 'N' sta per nucleo):

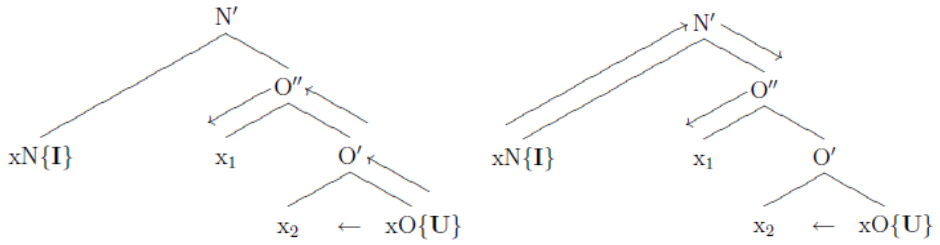
(12) a. [iː] b. [ɪ]



In (12a) la posizione $\times_1$ è m-comandata dal nucleo (annotato con l'elemento **I**) mentre in (12b) l'm-comando è esercitato dall'attacco.

In tal modo, l'informazione che nella ET standard è affidata all'elemento **H** è codificata nella struttura attraverso l'm-comando in GP 2.0. Questo è illustrato in (13) con la rappresentazione della sequenza vocale- consonante finale di parole come *rib* e *rip* (da Pöchtrager 2006: 71):

(13) a. [ɪːb] b. [ɪp]



A differenza della ET standard, in GP 2.0 le proprietà di modo di articolazione sono rappresentate in termini di struttura: una struttura più complessa corrisponde a più livelli di proiezione. Ad esempio un'occlusiva, rappresentata nella ET standard come **[ʔ H]** corrisponde a due livelli (O"), mentre una fricativa **[H]** corrisponde a un livello (O').

In questo sistema formale la lenizione può corrispondere a cancellazione di strati della struttura e quindi a una sua semplificazione; ad esempio la spirantizzazione può essere direttamente rappresentata come il taglio della proiezione più alta, che indica la rimozione dell'occlusione. Al contrario, questo sistema non rende possibile la rappresentazione di effetti di lenizione come quelli descritti in (9a) e (9b), che corrispondono nella rappresentazione standard alla perdita dell'elemento **H** ma non di **ʔ**. Una soluzione può essere trovata ammettendo che esiti di lenizione diversi siano in realtà il risultato di una riduzione di strutture segmentali diverse

in partenza. Questa ipotesi solleva però problemi di arbitrarietà della rappresentazione e potrebbe contrastare con il fatto che, in una stessa lingua, una consonante può andare incontro a fenomeni di lenizione diversi. Un caso di questo tipo si osserva in inglese britannico, in cui la glottalizzazione di /t/ con riduzione a [ʔ] si verifica in posizione preconsonantica e prepausale, come in *ge[ʔ] me* ma non in posizione prevocalica, come in *ge[t] it* (Harris 1994: 195-196).

4.2 Precedence-Free Phonology

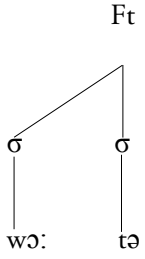
La teoria segmentale sviluppata nella PFPh non interviene sull'inventario degli elementi e mantiene il modello a sei unità definito nella ET standard. La revisione della GP classica individua come obiettivo di carattere generale la definizione del ruolo della fonologia all'interno della facoltà del linguaggio e la caratterizzazione del componente fonologico come parte integrante della facoltà del linguaggio in senso stretto (FLN nel senso di Hauser, Chomsky e Fitch 2002) (cfr. Nasukawa 2017a). Tale obiettivo è perseguito sulla base di alcuni assunti fondamentali della teoria. Il primo è che la relazione fra testa e dipendente nella struttura fonologica sia della stessa natura della relazione testa-complemento della sintassi. Nasukawa e Backely (2015), Nasukawa (2017b) sostengono che nei costituenti sintattici la testa apporti un contenuto informativo/semantico inferiore a quello portato dal complemento: nella combinazione sintattica la testa ha il compito di costruire la struttura attraverso la proiezione e di definire l'etichetta categoriale dell'intero costituente, mentre gli elementi non testa hanno un ruolo più importante nell'esprimere l'informazione lessicale. L'estensione di questa caratterizzazione alla fonologia produce un principio fondamentale della PFPh, secondo cui, nei diversi domini della struttura fonologica, le teste hanno il ruolo di costruire struttura attraverso la combinazione di tipo sintattico (X-barra) e la proiezione dell'etichetta, mentre i dipendenti contengono la maggior parte dell'informazione pertinente a ciascun livello della struttura; i dipendenti sono quindi "più grandi" e le teste sono "più piccole".

Il grado di informatività di ciascun elemento della struttura fonologica è stabilito in riferimento alla teoria del "*modulated carrier signal*" (Traunmüller 1994; Harris 2006, 2009). Secondo questa teoria, il parlato è costituito da un segnale acustico di sfondo (*carrier signal*) a cui si sovrappongono modulazioni che modificano più o meno intensamente il segnale stesso fino anche a interromperlo. Il *carrier signal* è un segnale acustico periodico con struttura formantica neutra (formanti equidistanti), prodotto con vibrazione glottidale senza articolazioni nel tratto sopralaringale, cioè una vocale neutra, il corrispettivo di **@** della ET classica. Le modulazioni sono dovute all'attività articolatoria che modifica la struttura formantica del segnale acustico, vi inserisce struttura aperiodica dovuta a restrizioni e ne provoca interruzioni totali dovute a occlusione. Le modulazioni trasportano il messaggio linguistico, mentre il *carrier* fa sì che il messaggio possa essere udito (Harris 2006). Nella PFPh, il dualismo *carrier signal/modulations* è reinterpretato attraverso la distinzione testa-dipendente nella struttura fonologica (Nasukawa 2017b). La testa costruisce la struttura dell'espressione linguistica, i dipendenti portano l'informazione pertinente a ciascun livello della struttura. Così, nella sillaba, è l'attacco che contiene modulazioni più forti rispetto al nucleo e quindi è "più grande" del nucleo: il nucleo è la testa.

In PFPh, l'argomento secondo cui la testa è "più piccola" e il dipendente è "più grande" assume il ruolo di un principio della rappresentazione fonologica, che ha l'esplicito scopo di rendere la fonologia e la sintassi uguali dal punto di vista della prominenza relativa (Nasukawa 2017b: 132). Pertanto, se nella sillaba è possibile mantenere la concezione tradizionale secondo cui la testa corrisponde al nucleo, lo stesso non vale per il piede. In un piede trocaico, la sillaba iniziale, che è generalmente considerata la testa in quanto sede dell'accento, è prominente e

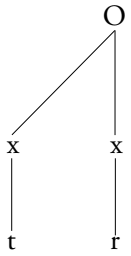
sede di modulazioni più forti (maggiore varietà vocalica, maggiore restrizione consonantica); la sillaba iniziale è quindi "più grande" dal punto di vista informativo. In PFPh questa caratteristica assegna alla sillaba accentata il ruolo di dipendente ed è la sillaba atona ad assumere il ruolo di testa: i trochei diventano in effetti giambi. In (14) è riportata la struttura del piede corrispondente all'inglese *water* [wɔːtə] 'acqua' (Nasukawa 2017b: 136):

(14)



Lo stesso procedimento è esteso a tutti i costituenti fonologici. Nella rima, la testa corrisponde al nucleo, perché la coda contiene modulazioni più forti. Ma all'interno di attacco e nucleo, è il secondo elemento, comunemente considerato dipendente, ad assumere il ruolo di testa in quanto "più piccolo" dal punto di vista informativo; questo è rappresentato in (15) per l'attacco complesso *tr* (Nasukawa 2017b: 140):

(15)



Non ci soffermiamo qui sulle implicazioni che questa argomentazione può avere sulla nozione stessa di testa, tanto in fonologia quanto in sintassi, e passiamo ad analizzare un altro aspetto essenziale di PFPh.

Il secondo assunto di questa teoria nega la concezione multilineare della rappresentazione fonologica, affermando che prosodia e melodia non appartengono a livelli separati. La rappresentazione fonologica fa riferimento solo ai primitivi melodici, cioè gli elementi (Backley e Nasukawa 2020), mentre i costituenti prosodici (i costituenti sillabici, il piede, la parola) sono proiezioni di teste corrispondenti a elementi. L'elemento corrispondente al *carrier signal,* che come tale è l'unità più povera dal punto di vista informativo, è la testa melodica che con le sue proiezioni genera la struttura fonologica per tutti i livelli di costituenza. L'elemento che assume questo ruolo fondamentale di costruttore della struttura fonologica è la testa di un nucleo e quindi appartiene al sottoinsieme degli elementi vocalici; questo elemento fondamentale resta lo stesso per tutte le espressioni di una data lingua.

Ogni lingua seleziona la testa melodica tra **A I U**. L'elemento selezionato si rivela nella forma della vocale di default, la vocale che, nelle appropriate condizioni, emerge in corrispondenza di un nucleo vuoto (tipicamente la vocale epentetica) (Nasukawa 2014). La vocale di default

corrisponde alla realizzazione degli elementi singolarmente interpretati, cioè **A** [ə] **I** [ɨ] **U** [ɯ]. Osservando la vocale di default si capisce qual è la testa melodica selezionata in una determinata lingua. Ad esempio (Nasukawa 2014):

(16)

	inglese	cilungu	giapponese
testa melodica	**A**	**I**	**U**
vocale di default	[ə]	[ɨ]	[ɯ]

La testa melodica proietta struttura del formato X-barra in cui la testa e le sue proiezioni costituiscono i tre livelli della tradizionale struttura prosodica nucleo, rima, sillaba (Backley e Nasukawa 2020). Le vocali di default in (16) hanno quindi questa struttura:

(17)

[ə]	[ɨ]	[ɯ]	
A”	I”	U”	‘sillaba’
|	|	|	
A’	I’	U’	‘rima’
|	|	|	
A	I	U	‘nucleo’

Gli elementi vocalici contenuti nella posizione di complemento determinano interamente la qualità della vocale, perché il dipendente è più forte e la testa diventa completamente silente. Ad esempio, in una lingua come l’inglese in cui la testa melodica è **A** (la vocale di default è infatti [ə]):

(18)

[ə]	[a]	[u]	[i]
A”	A”	A”	A”
|	|	|	|
A’	A’	A’	A’
|	/ \	/ \	/ \
A	A A	A U	A I

La posizione di complemento in (19) può essere arricchita ulteriormente attraverso struttura ricorsiva:

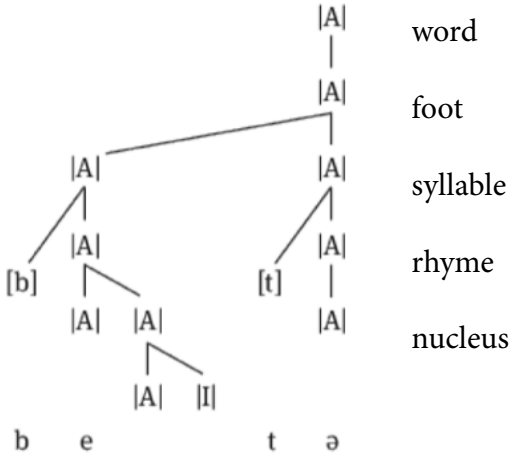
(19)

[æ]	[e]
A”	A”
|	|
A’	A’
/ \	/ \
A I	A A
/ \	/ \
I A	A I

Nella struttura ricorsiva la testa contribuisce alla qualità della vocale, in quanto fa comunque parte di un costituente dipendente, ma è recessiva rispetto al complemento dello stesso costituente, per cui $[\mathbf{A\ I}]_{\mathbf{A}}$= [e], $\mathbf{[I\ A]}_{\mathbf{I}}$= [æ].

Un'espressione melodica, o più propriamente un costituente, inserito nello specificatore, è interpretato come consonante. Questo livello di costituenda corrisponde alla sillaba e le proiezioni successive della testa **A** costruiscono struttura che corrisponde ai tradizionali costituenti prosodici (piede, parola). Ciò è illustrato nella rappresentazione in (20), tratta da Backley e Nasukawa (2020: 19), che corrisponde all'inglese *better* [betə]. Si tratta di una rappresentazione provvisoria e semplificata anche perché la struttura dell'attacco non è analizzata:

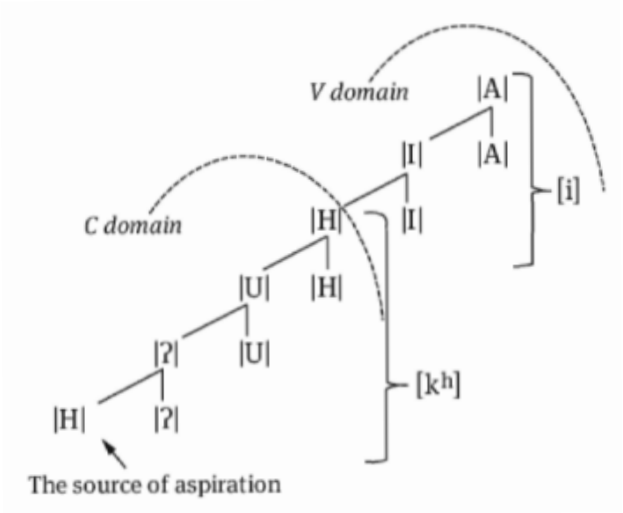
(20)



In tal modo, la consonante e la vocale che corrispondono a una sillaba CV sono unificati in una stessa struttura. Il concetto di Precedence-Free Phonology corrisponde al fatto che non ci sono relazioni di precedenza lineare tra le unità della struttura. Il fatto che la parte consonantica e la parte vocalica siano percepite come unità distinte non è codificato nella fonologia, ma dipende da restrizioni di livello fonetico: le proprietà consonantiche e quelle vocaliche sono fisiologicamente incompatibili e non possono essere realizzate simultaneamente (cfr. Nasukawa 2016; Onuma e Nasukawa 2020). Solo al livello fonetico, dunque, si produce la linearizzazione della struttura mostrata in (20), in cui i livelli più bassi della struttura (in notazioni alternative, i nodi dominati da A') sono interpretati simultaneamente, mentre i livelli più alti (in notazioni alternative Spec e A') sono linearizzati.

In questo modello le consonanti sono rappresentate come strutture costruite attraverso proiezioni verso il basso, con l'inserimento di ogni elemento ad ogni livello di proiezione. La rappresentazione in (21) si riferisce all'unità CV [k^hi] (da Backley e Nasukawa 2020: 31):

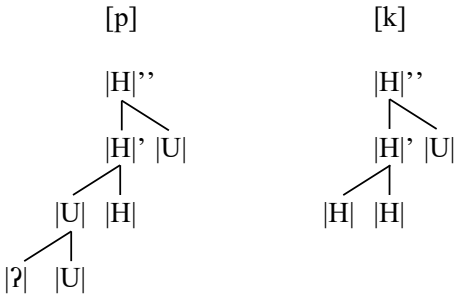
(21)



In (21) l'elemento **I**, che è dipendente nel dominio vocalico, diventa testa e si unisce a **H**, che caratterizza la consonante come ostruente. La testa **H** prende come complemento l'elemento di velarità **U**, che a sua volta diventa testa e prende come complemento **ʔ** che contrassegna la consonante come occlusiva. Infine, nella posizione di dipendente più bassa, e perciò più prominente, è inserito nuovamente **H**, che apporta la componente di aspirazione.

Le rappresentazioni proposte in lavori diversi nell'ambito di PFPh differiscono tra loro per alcuni aspetti riguardanti la forma dettagliata degli alberi, ma le caratteristiche del modello sono sufficientemente chiare riguardo alla struttura delle consonanti.[4] Ciascun elemento contenuto in un'espressione melodica di tipo consonantico è testa di una proiezione; inoltre, un elemento marcato come testa nel modello di Backley (2011), compare due volte nella struttura. Ad esempio l'elemento **U** che nella ET standard corrisponde al luogo velare e, se marcato come testa, al luogo labiale è presente una volta nella rappresentazione di [k]e due volte nella rappresentazione di [p]; questa configurazione è illustrata in (22), relativamente al giapponese (da Nasukawa 2016: 9):

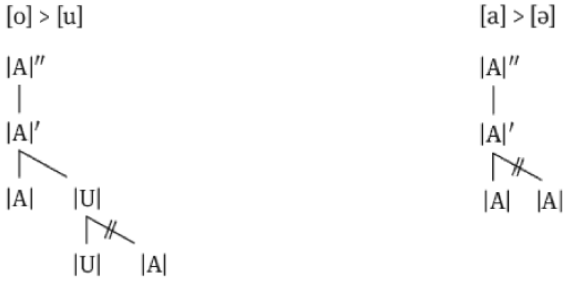
(22)



[4] Ad esempio, lo Spec, H" è regolarmente riempito nella rappresentazione delle consonanti in Nasukawa (2016), ma non in Backley e Nasukawa (2020). Per quanto di mia conoscenza, non è stata pubblicata una descrizione esaustiva e sistematica della struttura delle consonanti in PFPh, né del modo in cui vi si rappresentino i fenomeni di lenizione.

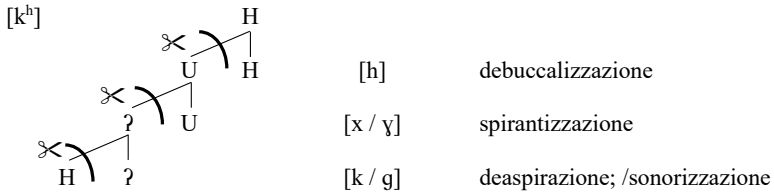
Data questa struttura delle consonanti e mantenendo la concezione dell'indebolimento segmentale come perdita di elementi propria della ET, i fenomeni di lenizione devono essere rappresentati come "potatura" di rami dell'albero. In tal modo, infatti, sono rappresentati i fenomeni di soppressione di elementi vocalici in Backley e Nasukawa (2020: 27):

(23)



Data dunque la struttura di una occlusiva velare sorda di una lingua come l'inglese già mostrata in (21), è possibile rappresentare alcuni esiti leniti tramite taglio di parti via via più grandi della struttura:[5]

(24)



In (24), dato che l'elemento **ʔ** è contenuto nel costituente dominato da **U**, l'esito della glottalizzazione completa con perdita di elemento di luogo di articolazione [k] → [ʔ] non è derivabile direttamente. L'unica soluzione è che la glottalizzazione corrisponda a una diversa disposizione gerarchica degli elementi, in cui **ʔ** sia dominato direttamente da **H**. Questa ipotesi, però, pone un problema di arbitrarietà della rappresentazione, dello stesso tipo quello rilevato precedentemente per GP 2.0.

Mantenendo un principio fondamentale della ET, secondo cui le categorie segmentali sono individuate sulla base del loro comportamento fonologico di cui la forma fonetica è solo un aspetto, è possibile affermare che quello che è descrittivamente un dato segmento, ad esempio [k], possa in realtà corrispondere a contenuti segmentali diversi o, in altri termini, a strutture

[5] La soppressione del **H** più basso produce una consonante neutra dal punto di vista delle proprietà laringali, che in lingue come l'inglese può manifestarsi come sonora in posizione intervocalica; lo stesso vale per l'esito fricativo (Onuma e Nasukawa 2020: 220, 221 e nota 3).

X-barra diverse, in cui uno stesso elemento può essere collocato a livelli diversi della gerarchia. Ricordiamo però che in una stessa lingua sono possibili esiti di lenizione differenziati per una stessa consonante, come si è visto nei paragrafi precedenti. Questa situazione richiederebbe l'ipotesi di strutture diverse per uno stesso oggetto fonologico, ponendo quindi lo stesso problema di arbitrarietà della rappresentazione già rilevato riguardo alla teoria GP 2.0.

5. Conclusioni

In questo articolo sono state presentate alcune riflessioni sulla capacità esplicativa di alcuni recenti modelli della ET in relazione ai fenomeni di lenizione. Si è osservato come, riguardo alla lenizione, questi modelli non offrano strumenti formali più potenti, o di più alto valore teorico rispetto alla ET delle origini, che ha fatto della rappresentazione dei fenomeni di riduzione un argomento cruciale a favore dell'ipotesi che i primitivi segmentali siano unità monovalenti e autonomamente pronunciabili, a differenza dei tratti binari della teoria generativa standard.

Le due teorie qui analizzate, però, mettono in luce un'altra questione, più generale e di grande interesse teorico, cioè la posizione della fonologia all'interno dell'architettura del linguaggio. Entrambi i modelli, infatti, perseguono l'obiettivo di rendere la fonologia più simile alla sintassi, avversando l'affermazione secondo cui la fonologia è diversa (cfr. Bromberger e Halle 1989). Il tema non è nuovo, e in particolare è una questione cruciale all'interno della GP, che ha fin dall'inizio individuato fra i propri fondamenti principi che escludono l'arbitrarietà dalle rappresentazioni fonologiche, impedendo, ad esempio, l'ordinamento di regole (cfr. ad esempio Kaye 2005).

In modelli come quelli qui analizzati, l'obiettivo di unificare sintassi e fonologia è perseguito attraverso l'affermazione che la regola fondamentale della sintassi, *Merge*, sia anche la regola della fonologia, la cui struttura sarebbe quindi di tipo "sintattico", e attraverso l'adozione di un formalismo caratterizzato da innumerevoli parallelismi con la sintassi, incluse specifiche configurazioni relazionali tra posizioni dell'albero (cfr. Nasukawa 2017a; Živanović e Pöchtrager 2010).

In generale, queste teorizzazioni non tengono conto di sviluppi della teoria sintattica che propongono revisioni profonde dei concetti stessi di "testa" e "proiezione" propri della teoria X-barra (Chomsky 2013, 2020; Chomsky, Gallego e Ott 2019). Più in particolare, nei modelli qui descritti l'operazione *Merge* all'interno dei segmenti crea una struttura gerarchica che è interpretata simultaneamente. I problemi di arbitrarietà descritti nei due paragrafi precedenti sono in effetti legati a questa caratteristica della rappresentazione segmentale, e proprio da questo punto di vista il parallelismo tra fonologia e sintassi si oscura. Infatti, se è vero che l'ordine delle parole può essere considerato come un effetto del sistema di esternalizzazione, proprio del componente morfofonologico e quindi estraneo alla FLN (cfr. Chomsky 2013, 2020; Chomsky, Gallego e Ott 2019), la linearizzazione della sintassi non può essere arbitraria. Il movimento sintattico che consiste in un'operazione di combinazione sintattica (*Internal Merge)*, e come tale appartiene alla computazione sintattica in senso stretto, è interpretato all'interfaccia semantica come una catena di occorrenze di un dato oggetto sintattico e all'interfaccia fonologica come un oggetto sintattico dislocato. Grazie alla dislocazione, dunque, l'operazione di *Merge* interno è riconoscibile e interpretabile semanticamente, e, per lo stesso motivo, apprendibile. Al contrario, l'ipotesi che un suono linguistico sia l'interpretazione fonetica simultanea di una struttura gerarchica complessa contenente più livelli di proiezione pone un cruciale problema di apprendibilità.

Nella prospettiva che la computazione fonologica faccia parte del nucleo della facoltà del linguaggio (FLN), è necessario che le spiegazioni delle proprietà fonologiche abbiano il carat-

tere di *genuine explanation* (Chomsky 2020) e possano quindi dimostrare che un determinato costrutto teorico attribuito alla Grammatica Universale possa essere acquisito dagli individui e sia compatibile con l'evoluzione della facoltà di linguaggio. Nella stessa prospettiva è anche necessario chiedersi se l'ipotesi che l'operazione *Merge* operi a tutti i livelli della rappresentazione fonologica, inclusa quella intrasegmentale, sia necessaria e opportuna, o se invece la FLN non possa contenere un sistema computazionale almeno in parte diverso per la parte della competenza linguistica che sta all'interfaccia con i sistemi di esternalizzazione.

References

Anderson, John M., and Charles Jones. 1974. "Three Theses Concerning Phonological Representations." *Journal of Linguistics* 10 (1): 1-26.

Anderson, Stephen R. 1976. "On the Description of Consonant Gradation in Fula." *Studies in African Linguistics* 7 (1): 93-136.

Backley, Phillip. 1993. "Coronal: the Undesirable Element." *Working Papers in Linguistics* 5: 301-323.

Backley, Phillip. 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh UP.

Backley, Phillip, and Kuniya Nasukawa. 2009. "Headship as Melodic Strength". In *Strength Relations in Phonology*, ed. by Kuniya Nasukawa, and Phillip Backley, 47-78. Berlin: De Gruyter.

Backley, Phillip, and Kuniya Nasukawa. 2020. "Recursion in Melodic-Prosodic Structure." In *Morpheme Internal Recursion in Phonology*, ed. by Kuniya Nasukawa, 11-35. Berlin: De Gruyter.

Bafile, Laura. 2019. "Vowel-Zero Alternations in Government Phonology and Strict CV Theory." *Studi e Saggi Linguistici* 57 (2): 83-113.

Bafile, Laura. 2020. "The Question of Overgeneration in Element Theory." In *Linguistic Variation: Structure and Interpretation*, ed. by Ludovico Franco, and Paolo Lorusso, 19-28. Berlin: De Gruyter.

Bromberger, Sylvain, and Morris Halle. 1989. "Why Phonology is Different." *Linguistic Inquiry* 20 (1): 51-70.

Carvalho, Joaquim Brandão de, Tobias Scheer, and Philippe Ségéral (eds). 2008. *Lenition and Fortition*. Berlin: De Gruyter.

Chomsky, Noam. 2013. "Problems of Projection." *Lingua* 130: 33-49.

Chomsky, Noam. 2020. *The UCLA Lectures* (April 29 - May 2, 2019). <https://ling.auf.net/lingbuzz/005485> (06/2021).

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

Chomsky, Noam, Ángel J. Gallego, and Dennis Ott. 2019. "Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges." *Catalan Journal of Linguistics - Special Issue*: 229-261.

Foley, James. 1970. "Phonological Distinctive Features." *Folia linguistica* 4(1-2): 87-92.

Goldsmith, John A. 1976. *Autosegmental Phonology*. PhD Dissertation, MIT.

Harris, John. 1994. *English Sound Structure*. Oxford-Cambridge, MA: Blackwell.

Harris, John. 1997. "Licensing Inheritance: an Integrated Theory of Neutralisation." *Phonology* 14 (3): 315-370.

Harris, John. 2006. "The Phonology of Being Understood: Further Arguments Against Sonority." *Lingua* 116 (10): 1483-1494.

Harris, John. 2009. "Why Final Obstruent Devoicing is Weakening." In *Strength Relations in Phonology*, ed. by Kuniya Nasukawa, and Phillip Backley, 9-46. Berlin: De Gruyter.

Harris John, and Geoff Lindsey. 1990. "Phonetic Interpretation in Generative Grammar." *UCL Working Papers in Linguistics* 2: 355-369.

Harris, John, and Geoff Lindsey. 1995. "The Elements of Phonological Representation." In *Frontiers of Phonology*, ed. by Jacques Durand, and Francis Katamba, 34-79. Harlow: Longman.

Harris, John, and Geoff Lindsey. 2000. "Vowel Pattern in Mind and Sounds." In *Phonological Knowledge: Conceptual and Empirical Issues*, ed. by Noel Burton-Roberts, Philip Carr, and Gerard Docherty, 185-205. Oxford, MS: Oxford UP.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. "The Faculty of Language: What Is It, Who Has It, and How did It Evolve?" *Science* 298: 1569-1579.
Hock, Hans. 1991. *Principles of Historical Linguistics.* Berlin: De Gruyter.
Honeybone, Patrick. 2008. "Lenition, Weakening and Consonantal Strength: Tracing Concepts Through The History of Phonology." In *Lenition and Fortition*, ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral, 9-92. Berlin: De Gruyter.
Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology.* New York, NY: Academic Press.
Hyman, Larry M. 1975. *Phonology. Theory and Analysis*. New York, NY: Holt, Rinehart and Winston.
Jensen, Sean. 1994. "Is ʔ an Element? Towards a Non-segmental Phonology." In *SOAS Working Papers in Linguistics & Phonetics* 4: 71-78.
Kaye, Jonathan. 1989. *Phonology: a Cognitive View*. Hillsdale, NJ: Lawrence Erlbaum.
Kaye, Jonathan. 2005. "GP, I'll Have to Put Your Flat Feet on the Ground." In *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 283-288. Berlin: De Gruyter.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1985. "The Internal Structure of Phonological Elements: a Theory of Charm and Government." *Phonology Yearbook* 2 (1): 305-328.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1990. "Constituent Structure and Government in Phonology." *Phonology* 7 (2): 193-231.
Kiparsky, Paul. 1982. *Explanation in Phonology*. Dordrecht: Foris.
Kirchner, Robert. 1988. *An Effort-based Approach to Consonant Lenition*. PhD dissertation, University of California.
Lass, Roger. 1984. *Phonology. An Introduction to Basic Concepts*. Cambridge: Cambridge UP.
Lass, Roger, and John M. Anderson. 1975. *Old English Phonology*. Cambridge: Cambridge UP.
Lavoie, Lisa M. 2001. *Consonant Strength: Phonological Patterns and Phonetic Manifestations*. New York, NY: Garland Publishing.
Loporcaro, Michele. 1988. *Grammatica storica del dialetto di Altamura*. Pisa: Giardini.
Nasukawa, Kuniya. 2005. *A Unified Approach to Nasality and Voicing*. Berlin: De Gruyter.
Nasukawa, Kuniya. 2014. "Features and Recursive Structure." *Nordlyd* 41 (1): 1-19.
Nasukawa, Kuniya. 2016. "A Precedence-free Approach to (De-) Palatalisation in Japanese." *Glossa: a Journal of General Linguistics* 1 (1): 1-21. doi: 10.5334/gjgl.26.
Nasukawa, Kuniya. 2017a. "Extending the Application of Merge to Elements in Phonological Representations". *Journal of the Phonetic Society of Japan* 21 (1): 59-70.
Nasukawa, Kuniya. 2017b. "The Phonetic Salience of Phonological Head-dependent Structure in a Modulated-carrier Model of Speech." In *Beyond Markedness in Formal Phonology*, ed. by Bridget D. Samuels, 121-152. Amsterdam: John Benjamins.
Nasukawa, Kuniya, and Phillip Backley. 2015. "Heads and Complements in Phonology: A Case of Role Reversal?" In *Phonological Studies* 18: 67-74.
Nocchi, Nadia, and Stephan Schmid. 2008. "Aspetti della lenizione in alcune varietà dell'italiano meridionale." In *La comunicazione parlata*, ed. by Massimo Pettorino, Antonella Giannini, Marianna Vallone, and Renata Savy, 109-136. Napoli: Liguori.
Onuma, Hitomi, and Kuniya Nasukawa. 2020. "Velar Softening Without Precedence Relations." In *Morpheme-internal Recursion in Phonology*, ed. by Kuniya Nasukawa, 207-236. Berlin: De Gruyter.
Paradis, Carole, and Jean-Franç Prunet. 1991. *The Special Status of Coronals: Internal and External Evidence*. In *Phonetics and Phonology 2.* San Diego, CA: Academic Press.
Pöchtrager, Markus A. 2006. *The Structure of Length*. PhD Dissertation, Universität Wien.
Pöchtrager, Markus A. 2020. "Recursion and GP 2.0." In *Morpheme-internal Recursion in Phonology*, ed. by Kuniya Nasukawa, 237-266. Berlin: De Gruyter.
Pöchtrager, Markus A., and Jonathan Kaye. 2013. "GP 2.0." In *SOAS Working Papers in Linguistics* 16: 51-64.
Roberts, Ian. 1997. *Comparative Syntax*. London - New York, NY: Arnold.
Schane, Sanford S. 1984. "The Fundamentals of Particle Phonology." *Phonology Yearbook* 1: 129-156.

Scheer, Tobias. 2004. *A Lateral Theory of Phonology*. Berlin: De Gruyter.
Scheer, Tobias, and Markéta Ziková. 2010. "The Coda Mirror v2." *Acta Linguistica Hungarica* 57 (4): 411-431.
Ségéral, Philippe, and Tobias Scheer. 2001. "La coda-miroir." *Bulletin de la Société de Linguistique de Paris* 96: 107-152.
Ségéral, Philippe, and Tobias Scheer. 2008. "The Coda Mirror, Stress and Positional Parameters." In *Lenition and Fortition*, ed. by Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral, 483-518. Berlin: De Gruyter.
Traunmüller, Hartmut. 1994. "Conventional, Biological, and Environmental Factors in Speech Communication: A Modulation Theory." *Phonetica* 51: 170-183.
Živanović, Sašo, and Markus A. Pöchtrager. 2010. "GP 2, and Putonghua too." *Acta Linguistica Hungarica* 57 (4): 357-380.

Semantica

Semantics

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: A. Peruzzi (2021) Noema fondato 2: schemi cinestetici e architettura categoriale della semantica. *Qulso* 7: pp. 207-231. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12009

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Noema fondato 2: schemi cinestetici e architettura categoriale della semantica*

Alberto Peruzzi
Università degli Studi di Firenze (<alberto.peruzzi@unifi.it>)

Abstract:

The paper describes the path of my research towards a *kinaesthetic theory of meaning* within which some of the main issues in 20th century's semantics and philosophy of language are approached from the perspective of a naturalized phenomenology centered on spatiality. The core issue lies in the *grounding* of cognitive structure, as expressed in supposedly "atomic" sentences, and the solution proposed rests on the existence of a finite set of basic schemata of objects-and-actions which, suitably combined, allow for the constitution of any possible act of reference. The steps which led to the *kinaesthetic theory* are presented here in retrospect, from the roots in Husserl's phenomenology to the essential debt toward cognitive grammar, particularly for what concerns the theory of metaphor. The emphasis is on one characteristic feature of the resulting approach, namely, the tools provided by category theory are argued to be indispensable in order to set up a model of the way human cognition and particularly logical architecture are grounded on basic schemata.

Keywords: *Category-theoretic Semantics, Cognitive Grammar, Kinaesthetic Schema, Metaphor, Naturalized Phenomenology*

1. "Ubicazione" della ricerca

La semantica si dice in molti modi. Purtroppo, la maggior parte di questi modi oscilla fra sottigliezze preteoriche ed eser-

[1] Questo articolo è una versione ampiamente rivista di un articolo, "Noema fondato: la prospettiva della semantica cognitiva", pubblicato nel volume *Fenomenologia applicata* (a cura di R. Lanfredini), Guerini, Milano, 2004, pp. 13-38. Il testo originario corrispondeva a una relazione dal titolo "Grounded Noema", tenuta nell'ambito del Convegno Internazionale *Mind and Language*, 15-18 ottobre 2003, presso l'Università di Bologna. Al pari del contenuto, nella presente versione la bibliografia originaria è rimasta pressoché invariata, per fotografare quello che allora era lo stato dell'arte. Le pochissime voci aggiunte testimoniano la scarsità di contributi significativi sul piano teorico dopo il 2004, a fronte di straordinari progressi della ricerca sul piano empirico, in linguistica così come nelle neuroscienze cognitive.

cizi teorici puramente formali. La teoria semantica sarà qui intesa come *anello di congiunzione* fra due vie: (A) la via della "mente incorporata" nell'ambito delle scienze cognitive e (B) la via categoriale ai fondamenti della matematica. L'idea di questo anello prese forma negli anni Ottanta all'interno delle mie prime ricerche su particolari problemi di semantica per risolvere i quali mi proponevo di sviluppare l'impostazione fenomenologica;[1] dunque, l'idea aveva dietro di sé una tradizione – e una letteratura – tanto ricca quanto ampiamente ignorata dagli scienziati cognitivi e da chi affrontava questioni fondazionali in termini assiomatici. Fino ad allora, la tradizione fenomenologica aveva a sua volta ignorato, e in seguito ha continuato a ignorare, le potenzialità di concreto sviluppo offerte da un simile anello.

Sia l'idea dell'*embodiment* sia l'idea di un fondamento categoriale collidono già, ciascuna per proprio conto, con i più diffusi orientamenti della filosofia del linguaggio del Novecento; inoltre, le due vie (A) e (B) si sono sviluppate l'una indipendentemente dall'altra e hanno lessici pressoché disgiunti. Perciò è facile capire che l'idea di uno stretto collegamento fra (A) e (B) sia apparsa a molti bizzarra e, sulla base di questa impressione, sia stata rapidamente liquidata.

Onde evitare quest'impressione, segnalerò (con taglio retrospettivo) alcuni punti di consonanza e dissonanza nei confronti della tradizione fenomenologica, senza però entrare in particolari di carattere storico-critico[2] e indicherò le modifiche da apportare all'analisi husserliana dei noemi al fine di elaborare un quadro teorico coerente che abbia potere esplicativo oltre che descrittivo. Il quadro risultante ha una serie di tratti in comune con l'impostazione data da George Lakoff alla semantica cognitiva.[3]

I contributi di Lakoff alla semantica e soprattutto la fondamentale concezione, elaborata con Mark Johnson, della metafora quale motore della cognizione, sono ormai riconosciuti come pietre miliari della linguistica post-chomskiana, mentre non sono riconosciuti nel loro carattere fenomenologico.[4] Fra i motivi di questo mancato riconoscimento non è trascurabile un dato: quando si è cercato di recuperare il nocciolo della lezione husserliana[5] liberandolo dai successivi stravolgimenti di tipo ermeneutico-esistenziale, il più delle volte s'è finito per fare filologia invece che filosofia; e quando si è voluto far valere tale lezione in merito a problemi di filosofia della mente, di epistemologia e di semantica, si sono trascurati tanto gli sviluppi sostanziali che nella seconda metà del Novecento ci sono stati nelle scienze cognitive quanto i non meno sostanziali sviluppi relativi alla logica e ai fondamenti della matematica.

L'approccio cognitivo alle metafore, fondato su schemi d'immagine e relative proiezioni è invece una tappa ineludibile per un'analisi fenomenologica dei caratteri che pertengono all'*embodiment* della mente, con particolare riferimento alle macrostrutture cognitive che si rivelano nel linguaggio. Quest'approccio ha esercitato una forte influenza sullo sviluppo delle mie idee e in scritti precedenti ho già testimoniato il debito che ho verso Lakoff, pur riconoscendo che sussistono punti di rilievo teorico e filosofico sui quali le nostre impostazioni prendono strade diverse.

[1] La prima formulazione di questo progetto si trova in Peruzzi (1981).

[2] Al riguardo, Peruzzi (1989).

[3] Entrambi diamo un peso decisivo al modo in cui la teoria semantica affronta la natura dei concetti matematici, ma ci sono significative differenze, alcune delle quali sono analizzate in Peruzzi (2017).

[4] Di carattere fenomenologico è infatti l'analisi del linguaggio emotivo e del linguaggio morale che è stata sviluppata da Lakoff, così come l'analisi che Lakoff e Johnson hanno fatto delle stesse dottrine filosofiche (cfr. Lakoff e Johnson 1980; Lakoff 1987; Lakoff e Johnson 1999). È curioso che verso tali analisi ci sia stata scarsa attenzione da parte di chi si richiama alla tradizione fenomenologica.

[5] Come suggerito nel capitolo conclusivo di Peruzzi (1988), che ancora conservava il lessico husserliano.

Entrambi concepiamo pensiero e linguaggio come essenzialmente *radicati* nella corporeità. Entrambi siamo convinti che la semantica può essere una scienza e che per esserlo non può conservare traccia di un inconsapevole realismo ingenuo e di nostalgie metafisiche, non può legittimare una filosofia ridotta ad esercizio di analisi logica e neppure può ridursi a elucubrazioni ermeneutiche. Entrambi, infine, pensiamo che il radicamento nella corporeità si realizzi solo liberandosi tanto da un'impostazione logico-computazionale quanto da una presunta fissazione univoca, referenziale, causale, del significato. Se liberarsene è necessario, non basta però a selezionare una determinata teoria del significato dotata di potere esplicativo. A tale riguardo è mia convinzione che il progresso di tale teoria dipenda essenzialmente dal reciproco arricchimento delle due vie (A) e (B) e, per illustrare le ragioni a sostegno di tale convinzione ripercorrerò le tappe attraverso le quali la convergenza con l'approccio cognitivo alla metafora si è concretizzata intorno alla spazialità *perisomatica*, quale aspetto decisivo della semantica cognitiva. Quanto invece alla divergenza da Lakoff, due sono i motivi principali.

Il primo motivo riguarda l'attacco sferrato da Lakoff e Johnson a ciò che chiamano "oggettivismo", nel quale si sovrappongono assunti tipici del razionalismo e del realismo (metafisico, ingenuo e scientifico). Questo attacco dovrebbe salvare una forma di naturalizzazione per dare un senso coerente alla corporeità, mentre gli argomenti che – sulla scia dell'Husserl della *Krisis* – sono addotti contro tale "oggettivismo" impediscono di salvarla. L'attacco risulta perciò un esempio di *overkilling*, anche se per altri aspetti la proposta di un "esperienzialismo" dal volto umano è condivisibile.

Il secondo motivo riguarda il rifiuto da parte di Lakoff – di nuovo in perfetta consonanza con Husserl – di dare alla semantica una forma propriamente teorica, perché ciò richiederebbe che fosse presentata assiomaticamente e formalizzata in termini matematici, il che produrrebbe, secondo Lakoff, un circolo vizioso, perché un simile assetto presuppone ciò che vorrebbe formalizzare. A tale rifiuto ho contrapposto una prospettiva incentrata sulla teoria matematica delle categorie, che permette di descrivere la *struttura* di una teoria in modo alquanto diverso da quello di derivazione logico-insiemistica e inoltre prospetta un modo di concepire i fondamenti della matematica alquanto diverso da quello "gerarchico", secondo cui l'albero di tutte le possibili teorie ha un'unica radice.

Per illustrare questa diversa prospettiva eviterò nella presente occasione dettagli di carattere tecnico, i quali richiederebbero familiarità tanto con la teoria delle categorie quanto con la teoria neurale del linguaggio (NTL) che è adottata da Lakoff. D'altra parte, una pur sintetica introduzione ai concetti di entrambe le teorie prenderebbe molte pagine, togliendo spazio a quelle osservazioni filosofiche che invece servono a vagliare l'eredità husserliana[6] e all'esposizione dei caratteri che definiscono l'anello di congiunzione fra (A) e (B).

S'impone, tuttavia, un'annotazione metodologica. Da tempo, fra coloro che si preoccupano di dare un formato propriamente teorico alle indagini di semantica, è opinione comune che non si possono ignorare gli strumenti forniti dalla logica, dalla teoria degli insiemi e dalla teoria della computabilità, in vista delle applicazioni che tali strumenti hanno in linguistica, psicologia e intelligenza artificiale. Ora, se un insieme di strumenti teorici non permette di risolvere i problemi, non si capisce per quale motivo richiederne l'impiego. Qui i problemi sono quelli che riguardano la *competenza semantica* e la morale tratta dai promotori di una concezione *embodied* del linguaggio è stata, appunto, che si può, anzi si deve, fare a meno di tali strumen-

[6] Per una definizione dei concetti fondamentali della teoria delle categorie e per un panorama dei suoi rapporti con la logica, vedi Peruzzi 1991; Lawvere e Schanuel 1997. Per una sintesi delle principali linee di ricerca della semantica cognitiva Johnson 1987; Lakoff 1987; Peruzzi 2000; Talmy 2000.

ti. Ma, se un apparato formale fallisce, non significa che non ce ne sia uno migliore: è così che la conoscenza è progredita, mentre l'accusa di circolarità finisce per ritorcersi contro ogni proposito di *spiegare* i fenomeni che attestano la competenza semantica. Nel caso in esame, la teoria delle categorie offre strumenti più raffinati, che nel loro insieme costituiscono la cornice più avanzata per impostare l'analisi del linguaggio e per delineare una fenomenologia della matematica. Dietro a quest'affermazione non c'è un atto di fede nei confronti di un miracoloso formalismo; c'è piuttosto l'esigenza che la filosofia del linguaggio e la filosofia della mente non si nutrano di paleo-ontologia. In particolare, l'efficacia della cornice categoriale si misura con il suo contributo a fornire un modello della stessa genesi dei significati.

Per brevità, nelle pagine che seguono saranno trascurati *importanti* aspetti relativi al processo di concettualizzazione (che in psicologia cognitiva si dice "categorizzazione")[7] e ai rapporti fra scienze cognitive ed epistemologia. Mi limiterò a dare solo un'idea di come si configura la semantica cognitiva, mettendo poi in rilievo che gli strumenti offerti dalla teoria delle categorie permettono di darle una sistemazione unitaria.

Una simile linea di ricerca è, come accennato all'inizio, tutt'altro che popolare. Qui, per incrinare lo scetticismo al riguardo, proverò a spiegarne motivazioni ed esiti in relazione ad alcuni problemi "classici" della semantica.[8] Non mi illudo di convincere della validità di questa linea chi non ha alcuna intenzione di lasciarsi convincere; confido, però, di rendere più difficile la liquidazione della semantica cognitiva da parte sia dei custodi della tradizione fenomenologica sia dei filosofi analitici.

2. Primi passi

L'idea di usare le categorie invece degli insiemi in una semantica modellistica per il linguaggio naturale fu avanzata nella mia tesi di laurea (1974). L'anno dopo presentai (nell'ambito del "Seminario del sabato", organizzato da Ettore Casari a Firenze) un'assiomatizzazione categoriale della teoria husserliana del noema, *entro la quale* si potevano inquadrare le ragioni addotte da Frege per distinguere tra significato (*Sinn*) e riferimento (*Bedeutung*). Il risultato era un'ontologia intensionale in cui la predicazione è analizzata non più in termini di $\in$ ma di una combinatoria algebrica di funzioni entro un unico universo di discorso: termini singolari e predicati esprimono entità dello stesso tipo (noematico), anche se poi si organizzano in strutture distinte, alle quali corrispondono distinte forme lessicali in molte lingue. In tale cornice, la funzione denotativa aveva come input non espressioni ma significati e così l'esistenza di espressioni non-sinonime coreferenti si traduceva nel fatto che tale funzione non è iniettiva. Non c'è dunque bisogno, come del resto non c'era stato per Frege, di passare per contesti modali o contesti di discorso indiretto per avvalorare la non-iniettività e, d'altra parte, non ce n'è bisogno neanche per precisare il legame fra intensione ed estensione.

All'interno di un simile quadro i difetti dell'usuale analisi logica della predicazione risultavano spia di una carente analisi fenomenologica: il processo costitutivo delle varie forme di significato non è un "di più", dispensabile da una rigorosa trattazione formale della predicazione (e dell'identità). Le conseguenze di questa carente analisi si manifestano proprio quando prendiamo in esame i concetti centrali della matematica, quindi non c'è bisogno di supporre alcuna frattura *di principio* tra linguaggio comune e linguaggi formalizzati e tanto meno fra logica matematica e logica filosofica.

[7] L'uso ampiamente diffuso del termine "categorizzazione" è improprio: ignora che da Aristotele in poi sono state elaborate varie teorie delle categorie, in nessuna delle quali il termine "categoria" e il termine "concetto" sono coestensivi.

[8] Per i relativi dettagli si veda, oltre ai testi indicati nella precedente nota 6, Peruzzi (2005).

Sebbene molti problemi affrontati *sperimentalmente* dalla scienza cognitiva fossero stati anticipati con la teoria husserliana dell'intenzionalità, questa teoria non ha consentito di precisarli e poi risolverli. Naturalmente, l'idea che le nuvole richiedano una descrizione nebulosa continua a far comodo ai filosofi, ma non fa progredire la filosofia. Viceversa, il tentativo, ad opera di alcuni filosofi d'ispirazione analitica, volto a recuperare le istanze husserliane traducendole in termini di "semantica a-mondi-possibili" si è lasciato sfuggire il senso, ben più radicale, di un'indagine *costitutiva* che in quanto tale va oltre l'assunto analitico secondo cui il pensiero è suscettibile di essere indagato solo in quel medium *lineare* che è la proposizione.

Da un lato, c'era la convinzione che la struttura del linguaggio può e dev'essere formalizzata mediante la logica e che la logica trova un'adeguata semantica nell'universo della teoria degli insiemi. D'altro lato, c'era l'idea che questa convinzione fosse utopistica e che non restasse altro che descrivere la molteplicità dell'uso contestuale del linguaggio. Anche chi, come Hilary Putnam (1999), ha poi parlato di una "triplice corda" fra mente corpo e mondo, riscoprendo la centralità della percezione, non lo ha fatto per impostare uno studio della strutturazione cognitiva dello spazio ma per rinverdire l'ecumenismo analitico.

Nel Novecento, questa polarità di atteggiamenti si è incrociata con un'altra: alla tradizione estensionale, fondata sul primato delle nozioni di riferimento e verità, si contrapponeva la tradizione intensionale, secondo cui il significato ha priorità sul riferimento e ne è (in parte) autonomo. Non è qui il caso di entrare nel merito della varietà teratologica di stratagemmi che sono stati tentati pur di legittimare quest'autonomia agli occhi dei fautori della tradizione estensionale. Gli stratagemmi più seri si limitavano a conservare un DOGMA: *le funzioni sono insiemi* (di *n*-ple ordinate). Quelli meno seri si tuffavano in un'ontologia di stampo medievale, del tutto inadeguata all'architettura del pensiero matematico. In entrambi i casi si prendeva per buono che una semantica rigorosa (diversa da quella sfuggente cui alludono i testi di semiotica) e non-esoterica (cioè, empiristica) può prender corpo esclusivamente in una cornice insiemistica, eventualmente arricchita con operatori modali.

Un danno collaterale del DOGMA era che quella particolare funzione denominata "interpretazione" (definibile induttivamente a partire da termini primitivi e predicati atomici) risulta una funzione totalmente libera e che molte "interpretazioni" empiricamente equivalenti sono sempre a disposizione. Di qui, non a caso, i ben noti argomenti di Hilary Putnam sui limiti della semantica modellistica e la morale antirealistica che se ne doveva trarre. Ma l'indefinibilità della verità (Teorema di Tarski) e del riferimento (Teorema di Löwenheim-Skolem) si possono addurre come prova definitiva dei limiti inerenti a qualsiasi semantica modellistica *se e solo se* ci si attiene al DOGMA.

Con le applicazioni dell'informatica allo studio della cognizione il problema del significato sembrava risolto: bastava considerare l'aspetto procedurale dei concetti-simboli. Purtroppo, la soluzione era truccata. Le procedure algoritmiche sono descritte come funzioni, dunque o le funzioni non sono ciò che i logici e i filosofi del linguaggio vogliono credere, o lo sono, e allora le procedure sono orpelli, dal momento che un algoritmo esecutore presuppone qualcosa che procedurale non è, perché le funzioni sono semplicemente sottoinsiemi di un prodotto cartesiano.

Un tanto elementare rilievo non esprime una critica alla semantica modellistica come tale ma agli strumenti con cui una tale semantica è stata generalmente sviluppata. Il rilievo è orientato alla dimensione procedurale come "costitutiva" del significato, perciò, che si accetti o si rifiuti la diagnosi di Richard Rorty sul destino ermeneutico della filosofia analitica, la cosiddetta "svolta linguistica" porta a un vicolo cieco: se il significato è dato dalle àncore referenziali a osservabili (extralinguistici), l'ancoraggio è o arbitrario o metafisico, e se il significato di un simbolo è dato da altri simboli (ora nel metalinguaggio), si produce un circolo vizioso oppure

un regresso all'infinito. In linguistica così come in matematica, c'è chi ha cercato di trasformare il circolo vizioso in uno virtuoso. Come? Pensando che l'intero sistema di simboli nelle loro mutue relazioni (la "struttura") definisce *implicitamente* il significato di ciascun simbolo. Chi l'ha pensato e chi ancora lo pensa suppongono tacitamente che una simile definizione implicita caratterizzi in maniera univoca il *definiendum.* Purtroppo, un classico teorema della logica matematica esclude tale supposizione per un'ampia classe di linguaggi.

La dimensione procedurale in questione è stata invece tematizzata nelle scienze cognitive. È così che si esce dal vicolo cieco? Questa via d'uscita è stata imboccata da molti, ma finché si suppone che il pensiero sia un superlinguaggio, il *remake* cognitivista dell'approccio analitico non funziona: nella migliore delle ipotesi, entro il modello della mente come elaboratore di informazioni, l'analisi del pensiero diventa analisi di un superlinguaggio di programmazione, capace di esprimere le più diverse specie di operazioni coinvolte nella competenza semantica. In tal modo potremmo anche apprezzare che l'astratta generalità della logica si riduca assumendo veste computazionale ma usciremmo dal vicolo cieco soltanto con un miracolo: l'intuizione diretta della sintassi del superlinguaggio. Un esito, questo, non certo nuovo, perché vi conduceva già l'idealismo di Berkeley (anche se il formato computazionale delle rappresentazioni ne restava ai margini) ed è alquanto curioso che, avendo poca dimestichezza con la filosofia, i promotori di modelli computazionali della mente abbiano riproposto idee che da più di due secoli sono state oggetto di numerose e ben argomentate obiezioni.

A scanso di equivoci: dietro al riconoscimento dei difetti di un simile *remake* non c'è lo sciovinismo specie-specifico dei filosofi che prima esclamano "Non siamo macchine!" e poi, eventualmente, cercano le prove di quanto hanno esclamato, riuscendo di solito a conquistarsi un facile consenso. Non c'è bisogno, però, di additare il ricorso alla *captatio benevolentiae* da parte di chi afferma l'irriducibilità della mente umana (e del linguaggio umano) a vili meccanismi, perché il difetto nel *remake* è più prosaico e, se non altro, più preciso: la teoria della computabilità, in virtù dell'equivalenza tra le sue diverse presentazioni, esige a sua volta una semantica ma se, per fissare le idee, ci riferiamo al lambda-calcolo, questa semantica non trova posto nell'universo standard degli insiemi, cioè, nella categoria i cui oggetti sono insiemi e i cui morfismi sono funzioni. Infatti, una volta ammesso che le operazioni di pensiero, aventi pensieri in input e pensieri in output, possono sempre essere oggettivate come pensieri e auto-applicate, le operazioni devono avere la forma di un insieme (per ipotesi) X di simboli mentali e di loro manipolazioni che sia chiuso rispetto all'esponenziazione, di modo che $X^X \cong X$; ma il solo insieme con questa proprietà è un singoletto ed è arduo proporre un simile X come modello perché la nostra mente è abitata da più pensieri.

Né, per ovviare in maniera empiricamente adeguata a quest'inconveniente, basta affidarsi a una teoria dei tipi come quella di Russell: infatti, una gerarchia di tipi con un unico tipo base (formato da individui, che l'empirista intende come "particolari") pregiudica la possibilità di trattare i fenomeni linguistici cui si indirizza la grammatica categoriale che ha preso avvio dalle husserliane *Logische Untersuchungen*. E neppure basta moltiplicare i tipi-base come in un generico lambda-calcolo con più *tipi*, anche se è indispensabile un linguaggio con più *tipi* (qui, non importa se di espressioni o di entità corrispondenti), perché solo con tale pluralità di tipi-base si riesce a dar conto di alcuni aspetti della concettualizzazione.[9] Il motivo per cui

[9] Se n'era già reso conto Richard Montague, pur non disponendo delle risorse fornite alla semantica dalla teoria delle categorie (cfr. i capitoli redatti da M. La Palme Reyes, J. Macnamara, A. Peruzzi e G. Reyes, in Macnamara e Reyes 1994). Tralascio qui di considerare la ripresa della via legata alla grammatica categoriale, così come si delinea a partire dal calcolo sintattico introdotto da Jim Lambek nel 1958. Per un confronto con la grammatica di Montague, si veda Carpenter (1998).

la "tipizzazione" non basta è che una mera pluralità di tipi non permette di selezionare alcuno *specifico* aggancio con la percezione. Per selezionarlo è necessario che la teoria semantica (I) espliciti la variazione (continua) delle entità cui ci riferiamo entro uno stesso dominio, combinata con la presenza di vincoli "morfici" che permettono la stabilità del riferimento allorché si passa da un dominio all'altro, e (II) espliciti il carattere costruttivo del ragionamento che si radica in *Gestalt* senso-motorie. Perché queste due condizioni trovino appropriata formulazione, è necessario (benché insufficiente) espandere il lambda-calcolo tipato giungendo a una logica d'ordine superiore in cui, oltre a disporre di più tipi-base, i quantificatori sono limitati a ciascun dato tipo τ, dunque si ha soltanto $\forall x{:}\ \tau\ \phi(x)$, invece di $\forall x\ \phi(x)$, e analogamente per $\exists$. In una simile cornice formale è possibile esprimere princìpi strutturali che precisino il nesso tra variabilità continua (I) e costruttività (II).

Intorno alla metà degli anni Ottanta mi resi conto che questa cornice c'era già: era stata messa a punto dalla trattazione categoriale della logica nei *topoi,* anche se nella sua piena generalità il nesso tra (I) e (II) restava ancora privo di quella tematizzazione fenomenologica,[10] con la quale sarebbe giunto a compimento il percorso iniziato vent'anni prima. Infatti, era stata proprio la ricerca del nesso appena menzionato a motivare le mie prime ricerche sulla semantica del linguaggio naturale: volevo mostrare che è possibile amalgamare una logica costruttiva con un'ontologia non statica e che si può simulare tale amalgama tenendo conto della prospettiva epistemica di un sistema/soggetto/agente *aperto* all'ambiente, prescindendo dal fatto che tale sistema fosse naturale o artificiale (dunque, senza alcun sciovinismo specie-specifico). Per riuscirci la tradizionale analisi del linguaggio non bastava perché, in vista di fornire un corrispettivo semantico della grammatica universale ipotizzata da Chomsky, cioè, un corrispettivo in grado di valere come teoria degli universali semantici, era indispensabile che l'interfaccia cognitiva fosse appunto tematizzata entro la cornice formale. Dunque, nel momento in cui si sostituiva il quadro strutturalistico con quello generativo nell'impostare una grammatica delle forme di significato accessibili agli umani, occorreva fare i conti con gli ostacoli incontrati dalla semantica generativa. Nei primi anni Settanta, tra i linguisti impegnati in quest'area c'era stato George Lakoff e, per testimoniare il debito che ho con lui, è opportuno un passo indietro.

Nel 1976 avevo frequentato un corso di Lakoff nell'ambito della *Summer School in Computational Linguistics* presso la Scuola Normale di Pisa. Grazie alle sue lezioni mi ero reso conto di quante questioni il mio iniziale progetto di ricerca stava trascurando. Tra gli argomenti del corso di Lakoff c'erano le difficoltà connesse all'analisi dell'*aspetto* verbale, le molteplici forme di costruttività logica compresenti nel linguaggio naturale (un'analisi che, sul piano della "struttura profonda", riguardava direttamente i processi di *costituzione* del significato), la presenza di vincoli gestaltici sull'ontologia del linguaggio naturale, e infine il carattere sfumato (*fuzzy*) della categorizzazione. Tutti questi argomenti erano affrontati non con lo spirito di chi vuol semplicemente trovare controesempi all'estensione del progetto chomskiano alla semantica, bensì con quello di chi intende dar conto di specifici fenomeni *cognitivi* soggiacenti al linguaggio. La morale era: nessuna delle teorie semantiche allora in circolazione, e fra esse anche la semantica generativa, è in grado di darne conto. C'era dunque bisogno di un approccio radicalmente diverso,[11] che a sua volta esigeva una diversa forma teorica.

[10] Limitatamente alla stabilità del riferimento di nomi propri e descrizioni, il requisito (I) è stato analizzato in termini di teoria dei gruppi, vedi Peruzzi (1987), mentre il senso fenomenologico del requisito (II) è stato messo in rilievo da Longo (2001) raccogliendo le preziose indicazioni di Hermann Weyl. Al riguardo è significativa la presa di distanza dell'ultimo Weyl da Husserl, come ricostruita da Bell (2004).

[11] Un primo spunto venne dalle ricerche di Jeffrey Gruber (cfr. Gruber 1976) poi riprese da Ray Jackendoff. Fu

Negli anni successivi la mia attenzione fu attratta dalla riabilitazione del ragionamento diagrammatico che si realizza con la teoria delle categorie e mi venne da pensare a un legame fra la spazialità implicita nell'uso di diagrammi e la mia precedente idea che le categorie siano la cornice adatta per trattare fenomeni intensionali: in teoria degli insiemi, le funzioni sono definite come insiemi, gli insiemi soddisfano al principio di estensionalità, e l'uso di diagrammi è solo un occasionale ausilio didattico, mentre in teoria delle categorie, grazie all'indipendenza del concetto di morfismo da quello di appartenenza, si può distinguere nettamente l'aspetto procedurale dalla composizione degli oggetti e la costruzione di diagrammi è appunto una procedura. Arrivai così a pensare che il concetto categoriale di funtore, originariamente motivato in relazione al nesso fra topologia e algebra e poi messo a frutto con i fasci, poteva servire anche a precisare l'ipotesi che la struttura logica del linguaggio sia una proiezione di *pattern* geometrici (o topologici) e a mettere a punto una corrispondente semantica. Le implicazioni di quest'ipotesi, che ha una duplice valenza, fenomenologica e fondazionale, erano molteplici e, per vagliarne le potenzialità in relazione a specifici problemi di analisi del linguaggio, invece di elaborare aggrovigliati argomenti filosofici, occorreva uno sguardo libero dai paletti del modello logico-linguistico imperante.

Di qui l'attenzione che in quegli anni rivolsi a un tema che era stato paradigmatico (da Russell in poi) per la filosofia analitica: le descrizioni definite. In alcuni lavori su tale tema argomentai che un'adeguata teoria delle descrizioni esigeva una cornice matematica in cui esprimere le condizioni di una stabile identificabilità delle entità descritte. Grazie all'incoraggiamento ricevuto da John Macnamara e Gonzalo Reyes sono poi riuscito a collegare tali condizioni a una più generale impostazione categoriale della semantica dei nomi propri e dei nomi comuni (esclusi i *mass nouns*).[12]

Incontrai nuovamente Lakoff in Austria, in occasione del convegno "Philosophy and the Cognitive Sciences" svoltosi a Kirchberg nel 1993. Lakoff tenne una relazione insieme a Mark Johnson nella quale si enfatizzava una terza "svolta" radicale nella filosofia del linguaggio, mentre il mio intervento verteva sul modo in cui gli universali semantici si possono caratterizzare nel linguaggio della teoria delle categorie. Conviene dunque ricordare brevemente le due "svolte" precedenti.

Nei primi decenni del Novecento c'era stata la prima svolta, la *svolta linguistica*, come poi denominata da Richard Rorty (riprendendo il titolo di un articolo di Moritz Schlick), che aveva dovuto il suo successo all'impiego della logica matematica. Ma, quando ne cominciarono a emergere i limiti nell'analisi del "linguaggio ordinario" (e curiosamente non se ne notarono i limiti nell'analisi del linguaggio matematico) la discussione restò confinata alla dimensione logico-linguistica oppure condusse a una nuova "casuistica", che rinunciava al compito teorico. In un modo o nell'altro, le questioni riguardanti la percezione, la causalità, il tempo e altri temi classici della filosofia recavano sempre il marchio di fabbrica: tutto quello che c'era da dire su X si limitava a un'analisi logica degli enunciati su X e ciò che non era legittimato come "dotato di senso" da tale analisi suggeriva i nuovi limiti della ragione.

La seconda svolta era stata la *svolta cognitiva*, che sfruttava i metodi dell'informatica e, anche nell'affrontare questioni semantiche, puntava tutto sull'architettura computazionale della mente. Di nuovo, nel momento in cui si riconoscevano i limiti dei modelli basati sull'elaborazione ricorsiva dell'informazione, si creava quel vuoto che permetteva di recuperare temi fenomenologici ed ermeneutici a uso e consumo dei più vari cantori del mistero, orgogliosi di ogni nuovo scacco dello "scientismo".

Massimo Moneglia a richiamare la mia attenzione su Jackendoff (1983), un'opera che ebbe un effetto dirompente sul quadro "modellistico" (*model-theoretic*) cui ancora facevo riferimento.

[12] In Macnamara e Reyes (1994) sono raccolti vari contributi che illustrano la ricchezza espressiva dell'impostazione categoriale per quanto riguarda l'analisi della predicazione nella semantica delle lingue naturali.

Con l'avvento della scienza cognitiva, l'*organon* dei filosofi analitici aveva subìto, sì, una mutazione, ma sotto le ceneri bruciava lo stesso fuoco: assimilando la mente a un insieme di programmi, la struttura logica del linguaggio si trasformava in un sistema di regole che definivano il Linguaggio del Pensiero. La mutazione trascinava con sé i limiti esplicativi di una semantica formale, elaborata con gli esclusivi strumenti della logica – e ora, più specificamente, della teoria della ricorsività – e proposta come asse portante di una filosofia rigorosa. La semantica formale, o "modellistica", andava incontro alla *fallacia simbolica*: o il circolo vizioso o il regresso all'infinito, cui ho già fatto riferimento. La semantica procedurale si ritrovava lo stesso ostacolo. D'altro lato, l'ancoraggio referenziale a qualcosa di non linguistico era etichettabile come un *hic sunt leones*.

Eravamo ancora una volta di fronte allo stesso guaio che Husserl aveva avuto il coraggio di affrontare in maniera decisiva: da quale porta il significato entra nella rete simbolico-formale? E, come ai tempi di Husserl, le strategie più seguite si ispiravano a una dilatazione della rete, fino al punto di dire che non c'è bisogno di alcuna porta perché, al pari di tutto ciò cui ci riferiamo, noi stessi *siamo* costrutti simbolici, oppure decretavano un solco incolmabile tra il Formale e il Vitale. Alcuni facevano di necessità virtù, specie nell'ambito della corrente logicista negli studi di intelligenza artificiale: dall'ambito celeste dei fondamenti il formalismo veniva trasportato all'ambito sublunare, quello del quotidiano commercio con la comprensione di significati e in questo modo il dominio di ciò che Husserl aveva chiamato un "significato-di-gioco" finiva per ricoprire l'intero mondo del senso comune.[13] Altri ne erano indotti a ricavarne il su menzionato *hic sunt leones,* dunque si riproponeva l'alternativa emersa decenni prima all'interno della filosofia analitica, ora però riformulata in termini della mente come sistema di programmi per la manipolazione di simboli.[14]

La terza svolta cui Lakoff e Johnson si riferivano era dovuta all'idea della *embodied mind*. In base a quest'idea anche le strutture più astratte del pensiero sono inscindibili dalla corporeità. La mia relazione a Kirchberg segnalava invece il non meno radicale cambiamento avvenuto nella semantica dei linguaggi formali e di programmazione con il passaggio da una concezione insiemistica a una categoriale. Fu immediato pensare che le due linee di ricerca, partite da problematiche distinte, potessero convergere.

3. Semantica cognitiva e naturalizzazione

L'idea della *embodied mind* che prende campo negli anni Ottanta non produce soltanto un cambiamento di prospettiva ma pervade gli studi su specifici fenomeni linguistici, portando a rintracciare nella stessa struttura delle frasi una spia dei modi fondamentali di integrare percezione visiva, sistema motorio, cinestesi e immaginazione. Non si tratta tanto di uno studio interdisciplinare quanto di un nuovo quadro concettuale che si configura direttamente attraverso paradigmatici *case studies* che permettono di identificare una lista di schemi concettuali prima ritenuti secondari nell'analisi logica.

È così che la *semantica cognitiva* si è precisata attraverso i lavori di George Lakoff e Mark Johnson sulla metafora, di Leonard Talmy sulla concettualizzazione dello spazio, di Ron Langacker sulla struttura enunciativa associata alla coppia *trajector/landmark* (Langacker 1987, 1991). Un punto di riferimento era la lezione di Eleanor Rosch sulla funzione delle *specie di livello basico* e sulla

[13] Al riguardo è significativa la rassegna offerta da Smith (1992).

[14] Winograd e Flores (1986). È per un'analisi critica delle presunte conseguenze ermeneutiche che Winograd e Flores traggono dal fallimento dei modelli computazionali dell'intenzionalità, si veda Peruzzi (1996).

finestra cognitiva entro la quale si costituiscono i *prototipi* concettuali. Inoltre, si recuperavano le ricerche pionieristiche di Charles Fillmore (1968, 1982, 1985) sui *frame*, con l'individuazione dei ruoli "tematici" fondamentali, estromessi fino ad allora dalla semantica (modellistica), e si profilava un'integrazione dell'approccio guidato dall'idea della *embodied mind* sia con le neuroscienze sia con l'approccio biologico-sistemico, proprio della concezione "enattiva" dell'intenzionalità.[15]

A Lakoff e Johnson va riconosciuto il merito di aver messo a punto un'estesa fenomenologia delle strutture soggiacenti alla metafora (preposizionale e verbale, prima che nominale) in modo da correlare alle radici corporee del significato una serie di *image schemas,* di cui vengono documentate le innumerevoli proiezioni metaforiche a dominî relativi a entità non fisiche, giungendo alla tesi che la visione del mondo che caratterizza ciascuna cultura si definisce attraverso una specifica selezione di *image schemas* e loro proiezioni. Emergono così due tipi di *pattern*: (a) *pattern* macrodinamici (come per esempio lo schema della verticalità che determina l'orientazione su/giù), grazie ai quali si organizza l'effettiva esperienza, nell'ambiente naturale e sociale, e si orienta l'immaginazione, (b) *pattern* di trasferimento-di-significato da un dominio all'altro e, più in generale, dal Concreto all'Astratto.

Si è così delineato il progetto di una "grammatica cognitiva", in cui i modelli metaforici hanno importanza primaria. La vasta gamma di questi modelli è stata addotta a sostegno dell'apertura *indeterminata* del pensiero umano, in contrasto con la rigida griglia della semantica d'ispirazione logica che di per sé lasciava ugualmente indeterminata l'interpretazione, ma una volta congiunta con le assunzioni del modello computazionale della mente e con le assunzioni realistiche implicite nel concepire il pensiero quale "specchio della natura", diventava il bersaglio da colpire. Tuttavia, chi adotta la prospettiva di una semantica cognitiva ancorata alla corporeità non è tenuto a sottoscrivere quest'apertura indeterminata: il rifiuto della purezza formale della logica e della diretta specularità fra proposizioni e fatti[16] non prelude infatti al rifiuto di ogni vincolo, ma piuttosto sollecita l'individuazione di specifici vincoli di ordine *naturale*.

L'enfasi sulla liberazione dalla rigidità della forma logica finisce dunque per essere fuorviante perché collide con l'impegno esplicativo assunto da chi riconosce i due tipi di *pattern* (a) e (b) come pervasivi di ogni possibile significato comprensibile dagli esseri umani, dunque anche l'efficacia che la forma logica ha per la razionalità che si esprime nel pensiero scientifico ha bisogno di essere spiegata. Il riconoscimento di quest'impegno esplicativo porta invece a una *Tesi* ben più sostanziosa di quel generico e confuso richiamo alla corporeità che ha fatto proseliti in campo fenomenologico:

> (*Tesi*) *Il pensiero consiste in un sistematico lifting di struttura semantica da una base ridotta, incentrata su un determinato insieme finito di* gestalt *figurali, ovvero, schemi posizionali di oggetti nello spazio e schemi basilari di azioni,*[17] *che, combinate in conformità a parametri contestuali, coprono tutto l'arco della cognizione umana.*

Questa tesi si coniuga con la geometrizzazione della logica in ambito categoriale, ma è poco condivisa all'interno della semantica cognitiva e le divergenze al riguardo si affiancano ad altre divergenze sul piano euristico-metodologico, già attestate del resto dai numerosi modi di impostare l'analisi delle metafore: per fare un solo esempio, la "teoria" del *blending*, avanzata

[15] Purché depurata da un certo olismo misticheggiante, la concezione enattiva, avanzata da Varela, Thompson e Rosch (1991), fornisce un antidoto a quel *semantic ascent* (per usare le parole di Quine) che condannava la filosofia al piano metalinguistico.

[16] Alcuni di questi vincoli sono forniti in Peruzzi (1996a).

[17] Per gli "*x*-schemi" (ove *x* sta per '*executing*'), si veda l'accurata indagine sviluppata in Bailey (1997) e la panoramica offerta da Freeman e Nùñez (1999).

da Gilles Fauconnier, ha caratteri diversi dalla "teoria" di Lakoff.[18] Inoltre, anche se il riorientamento degli studi linguistici sulla base dell'*embodiment* si è sviluppato insieme ai modelli a reti neurali, il tipo di connessionismo che interessa la semantica cognitiva è *strutturato.*[19] In base ai modelli PDP (*Partial Distributed Processing*), le "rappresentazioni" mentali, alle quali abbiamo accesso cosciente, sono il risultato di una rete massicciamente parallelistica di unità sub-logiche; dunque, quelle che alla riflessione appaiono come regole non sarebbero altro che epifenomeni di un'auto-organizzazione priva di specifiche regole. Senonché… l'idea che mediante *backpropagation* si effettui quella rapida selezione dei *pattern* che è dato osservare nelle prime fasi dello sviluppo mentale va incontro alle stesse obiezioni che Chomsky mosse al comportamentismo. Una significativa novità sta piuttosto nel fatto che gli effetti di soglia richiesti da tali modelli sono descrivibili con l'Analisi matematica (il Calcolo) e non per mezzo dell'analisi logica. Dunque, già all'interno di questi modelli si ammette la priorità del continuo sul discreto, e della geometrizzazione sull'aritmetizzazione del contenuto informativo nel raccordare linguaggio e sistemi percettivi; a quest'ammissione, però, non fa seguito una teoria geometrica del significato, come invece previsto dalla *Tesi*.

Con i modelli elaborati all'interno della NTL, così come delineata da Jerome Feldman, il quadro è cambiato ed è diventato possibile fronteggiare le critiche mosse al connessionismo, per esempio le famose obiezioni di Jerry Fodor e Zenon Pylyshyn. Se, partendo da un modello a reti neurali, si riesca a far fronte pienamente a tali critiche, è però difficile da valutare. Per farvi fronte possiamo arricchire il modello? Certo, ma, se ciò che aggiungiamo è di diversa natura, stiamo truccando le carte. Per esempio: il tempo (interno) delle reti connessioniste è tanto discrèto quanto lo è quello di una macchina di Turing, con difficoltà di sincronizzazione fra i diversi processi "sub-cognitivi" coinvolti, la cui integrazione è indispensabile se si vuol giungere a un'analisi soddisfacente almeno dei più elementari fenomeni semantici, ma al crescere della complessità sintattica le difficoltà aumentano e, se tale integrazione è pilotata dagli attrattori di un sistema dinamico, il quadro teorico non è più quello in cui i modelli a reti neurali erano stati pensati.

Ciò non toglie che i modelli connessionisti siano stati un passo significativo per ricostruire la gerarchia di processi che rende possibile l'accesso al significato di proposizioni elementari. Potremmo allora dire che analisi (logica) e Analisi (matematica) sono entrambe necessarie anche se non sufficienti a chiarire come tale accesso sia possibile, dovendo essere integrate con una serie di vincoli gestaltici, statici e dinamici. Ora, poiché una *gestalt* resta pur sempre una forma geometrica (in senso lato), ai fini di quest'integrazione *occorre* abbandonare il riduzionismo classico (dalla geometria all'Analisi e da questa alla teoria degli insiemi) e intendere la spazialità cognitiva come ambito primario a partire dal quale emerge la manipolazione numerico-simbolica, collocando in tale ambito le radici intuitive (corporee) di ogni significato. Di conseguenza, la geometria (in senso lato) si presenta come teoria generale della spazialità primaria e non è più formalizzabile nei termini della topologia degli "insiemi di punti".

[18] Fauconnier (1997) e Fauconnier e Turner (1998). Le virgolette al termine "teoria" sono d'obbligo, perché in entrambi i casi, come del resto in ogni altro progetto finora avanzato sul tema, si è ben lungi dalla formulazione di una teoria. Anzi, come già accennato, lo stesso proposito di formularla è stato contestato in quanto reo di circolarità. Per una replica, si veda la discussione critica sul concetto di schema, condotta in Peruzzi (1999).

[19] Le ragioni della differenza tra architettura globalmente distribuita e il connessionismo "strutturato" (*constrained*) della NTL richiedono dettagli tecnici la cui descrizione non è consona al tono discorsivo del presente articolo. Mi limito a rimandare a Feldman, Narayanan (2004) per un'esposizione introduttiva e a Regier (1996) per una trattazione focalizzata sul modo in cui la spazialità si organizza nel linguaggio.

Insomma, il recupero della spazialità e dei suoi *pattern* gestaltici va di pari passo con la naturalizzazione del significato ma si tratta di una naturalizzazione non riduzionistica, e proprio con ciò si manifesta uno dei motivi di divergenza dalle idee correnti circa lo status dei modelli metaforici, perché per molti degli studiosi di grammatica cognitiva il problema di rivedere la definizione insiemistica del concetto di spazio non si pone neppure, in quanto la geometria e, più in generale, la topologia sfruttano pur sempre costrutti metaforici, nessuno dei quali ha quindi titolo a una priorità cognitiva, mentre le osservazioni precedenti attribuiscono valenza fondazionale a tale priorità.

Tra la linea teorica che si esprime nella precedente *Tesi* e le linee più seguite nella semantica cognitiva c'è una ulteriore divergenza, concernente il modo di raccordare la semantica con i risultati di recenti indagini in ambito neuroscientifico.[20] Fermo restando che la NTL si differenzia dal connessionismo PDP, la NTL non è sufficiente a ricostruire pienamente una semantica *embodied* perché la combinatoria dei *pattern* gestaltici esige un'integrazione selettiva e stabile fra macro-componenti la cui "meccanica" ha a che fare con l'ambiente fisico. Per dar conto di questa integrazione è perciò indispensabile l'impiego di modelli a cicli di percezione-azione che sono descrivibili in termini di sistemi dinamici.[21] Ma la più generale cornice matematica in cui inquadrare la teoria dei sistemi dinamici è di nuovo quella categoriale, quindi anche sotto questo profilo la linea teorica che porta alla *Tesi* permette di apprezzare direttamente gli strumenti offerti dalla teoria delle categorie negli studi cognitivi,[22] a partire dall'idea di fondo, che la semantica cinestetica fa propria, secondo la quale ogni metafora è analizzabile in termini di un fibrato $p: X \longrightarrow B$ di cui ciascuna sezione locale s (esemplificante un lifting) associa a una data parte selezionata U di pattern di oggetti-e-azioni nel dominio letterale (B) una parte corrispondente in un dominio traslato (un fattore di X). La mappa p esprime il *radicamento* (che in inglese si indica solitamente come *grounding*) dei noemi e, trattando X e B come categorie, s si configura come un funtore di *metaforizzazione*, che trasferisce alcune proprietà relative alla parte U assunta come sorgente di significato letterale. Se nel quadro dinamicista ipotizziamo l'esistenza di attrattori cinestetici, la fecondità dell'impostazione categoriale si misura con la possibilità di caratterizzare tali attrattori mediante funtori aggiunti, generalizzando quella che nel 2000 ho chiamato "aggiunzione aurea" fra prototipi e specie di livello basico.

La via originariamente seguita negli studi di grammatica cognitiva, e in particolare da Lakoff per quanto riguarda le metafore, si contrapponeva alla semantica formale per ragioni propriamente linguistiche, non perché l'alternativa fosse motivata da sviluppi nelle neuroscienze. A maggior ragione, il fatto che questa via abbia trovato conferme in ambito neuroscientifico grazie ai lavori di Jerome Feldman, Terry Regier e poi soprattutto di Srini Narayanan è stato di straordinaria importanza per il progresso degli studi di grammatica cognitiva. Pur muovendosi nel solco di questi studi, la semantica cinestetica è stata invece motivata principalmente dall'esigenza di fornire una modellizzazione formale adeguata alla semantica e non può ancora vantare conferme analoghe.[23]

[20] A questo proposito, un fondamentale punto di riferimento è stato offerto dalla scoperta dei neuroni-specchio a opera del gruppo di ricerca coordinato da Giacomo Rizzolatti. Per un importante sviluppo successivo, si veda Gallese (2003).

[21] È la via dell'auto-organizzazione dinamica, definita in Kelso (1995) e associata allo slogan della *mente come movimento.* Al riguardo, si vedano i contributi raccolti in Port e van Gelder (1995) e in Peruzzi (2004).

[22] Peruzzi (2000) § 7 e, per un quadro più generale, Peruzzi 2004a. In precedenza, Goguen (1999) aveva segnalato l'interesse applicativo del linguaggio categoriale in ambito semiotico.

[23] Solo recentemente l'impostazione categoriale della teoria cinestetica ha trovato una prima, inaspettata, applicazione nelle indagini sulla struttura dell'ippocampo: Mancin (2020).

Non è escluso che queste due linee di ricerche trovino punti d'incontro, ma un ostacolo è costituito dalla rapida lezione filosofica che nel 1980 Lakoff e Johnson avevano inteso trarre dalla loro fenomenologia dei modelli metaforici: il nucleo di tale lezione starebbe nel riconoscimento che la dottrina, al contempo ingenua e iperformalizzata, da essi indicata come *oggettivismo,* è risultata fallimentare, confermando così la critica mossa da Husserl nella *Krisis* a una visione filosofica da lui indicata con lo stesso nome. La ragione per cui questa critica è d'ostacolo sta nel fatto che, anche ammettendo il fallimento dell'oggettivismo, resta del tutto oscuro in quale modo si possa simultaneamente affermare, con coerenza, il carattere *relativo-alla-cultura* di tutte le metafore e il carattere *embodied* della mente. In altre parole: se il modello delle scienze naturali è oggettivistico, e dunque viziato dal realismo, sia esso ingenuo o scientifico, è impossibile naturalizzare la semantica, ma allora il rimando alla corporeità è a sua volta carico di metafore essenzialmente culturali, quindi non può avere la generalità che gli è stata ascritta. In particolare, saremmo costretti a dire che la teoria neurale del linguaggio, pur avendo potere esplicativo, non ha potere selettivo (la selezione essendo esclusivamente culturale), come dire che il cervello è qualcosa di naturale ma totalmente plastico – ma allora le conferme in ambito neuroscientifico ne sarebbero fortemente ridimensionate.

Nell'idea che la semantica formale comporti necessariamente un impegno verso il realismo metafisico, Lakoff si è servito di due argomenti introdotti da Hilary Putnam contro tale realismo, mentre Johnson, in polemica con la filosofia analitica, ha raccolto i suggerimenti di Paul Ricoeur (di cui è stato allievo) per dare un taglio ermeneutico alla filosofia del linguaggio, liberandosi della cortina retorica di cui sul suolo europeo tipicamente si sono ammantati i testi degli ermeneuti. Incidentalmente: è fonte di stupore il fatto che proprio coloro che hanno voluto mettere in risalto la centralità dell'*interpretare* non si siano misurati con i problemi inerenti alla definizione induttiva, a partire da Tarski, del concetto di *interpretazione*.

Il guaio è che se, accantonati gli inutili scrupoli della semantica formale, ci atteniamo a un simile corredo filosofico della grammatica cognitiva, la stessa corporeità oscilla tra la plasticità dei circuiti neurali e la plasticità antropologica del significato. Nel primo caso la corporeità è fisica ma non selettiva, nel secondo è selettiva perché costruita, carica-di-cultura e relativa a modelli socialmente condivisi, ma per ciò stesso non ha più niente di fisico: è una corporeità metaforica. Come queste due nozioni di corporeità stiano insieme non è chiaro. La naturalizzazione che invece ho proposto senza rinunciare a quegli inutili scrupoli è *emergente* ma resta ancorata alle proprietà del macrospazio rilevanti dal punto di vista biologico e, ancor prima dal punto di vista fisico, perché hanno a che fare con la dinamica dei corpi materiali, e al riguardo il desiderato ecumenismo interculturale della lezione che si doveva trarre dalla fenomenologia dei modelli metaforici rivela i suoi limiti. La fisica non è quella monolitica fabbrica di modelli riduzionisti che molti filosofi vogliono credere per suffragare il loro diritto a ignorarla nonché la speranza di poter continuare a parlare di conoscenza e di realtà nel linguaggio comune, come in prevalenza hanno fatto dal V secolo a.C.. Qualunque sia la teoria semantica che prediligiamo, essa offre pur sempre un'analisi, condotta da qualcosa che fa parte del cosmo, di qualcosa che fa parte del cosmo, e la cosmologia non può prescindere dalla fisica.

A Kirchberg, dalle conversazioni con Lakoff emersero numerosi motivi di accordo così come alcune mie riserve, che qui ho sommariamente ricordato, circa l'attacco all'*oggettivismo* e, più specificamente, circa la coerenza di quest'attacco con il carattere auspicabilmente oggettivo della semantica cognitiva.[24] Non si trattava, a guardar bene, di una novità, perché la stessa

[24] La rilevanza di queste riserve nei confronti del disegno husserliano era già identificabile attraverso argomentazioni dall'apparenza innocua come quelle esposte in Peruzzi (1988) e, per ulteriori dettagli, Peruzzi (1996).

questione si era già posta per l'oggettività dell'analisi kantiana delle strutture della soggettività e c'era solo da sperare che con la grammatica cognitiva non si volesse riproporre quell'idealismo *soft* che è tornato di moda come congiunzione fra il prisma indefinitamente aperto degli a priori, ora intesi come convenzioni sociolinguistiche, e un non meno irenico pragmatismo. Benché i motivi di disaccordo già segnalati si ripropongano a proposito della descrizione che George Lakoff e Raphael Nùñez hanno dato del *romance of mathematics*, come risultante dalla congiunzione fra platonismo e naturalismo,[25] i punti in comune non mancano.

Al pari di Lakoff, nella relazione che tenni al congresso di Kirchberg nel 1993 esordivo indicando il vicolo cieco cui hanno condotto prima la *svolta linguistica* e poi la *svolta cognitiva*.[26] Infatti, finché la semantica resta confinata alla dimensione puramente formale, prima logica e poi computazionale, resta inspiegata l'interfaccia cinestetica attraverso cui si realizza l'accesso al significato. Gran parte della linguistica, al pari della "filosofia del linguaggio ordinario", nonché la fenomenologia vestita da psicologia del mistero, si sono trovate d'accordo nel reagire all'impiego di astratti formalismi contrapponendo la più a buon mercato delle soluzioni: spalmare la contestualità pragmatica su tutto e allo stesso tempo rivendicare l'autotrascendenza della mente da ogni contesto. Questa reazione non porta da nessuna parte, perché dà per scontata l'esistenza di menti e le condizioni che la rendono possibile. Per esempio, chi si richiama a Wittgenstein nella denuncia di residui platonici nell'analisi del linguaggio si ferma tipicamente alle "somiglianze di famiglia", senza esaminare tali condizioni, quasi che si potessero scindere da quelle che rendono possibile la flessibilità, tutt'altro che indeterminata, dell'uso linguistico nel macro-spazio dell'ambiente terrestre. Se invece ci risolviamo a considerare tali condizioni, non possiamo ignorare le scienze naturali e allora ci troviamo impegnati a individuare i princìpi che, definendo la finestra dei significati possibili per esemplari della nostra specie, fungono da *universali* (schematici) per l'interfaccia cinestetica relativa alla spazialità che ci si offre all'ordine di grandezza del nostro corpo.

A tali considerazioni d'esordio faceva seguito l'enunciazione della linea seguendo la quale si possono identificare tali princìpi, dando vita a una fenomenologia della cognizione semantica.[27] Non meno esplicito era il carattere peculiare di una concezione naturalistica che ospita un'analisi delle nozioni di verità e di riferimento senza dover sottoscrivere il realismo metafisico. Nel 1994 avevo dato il nome di "naturalismo intrecciato" a questa concezione in cui raccordare epistemologia e semantica, segnalando che la *naturalizzazione* risultante è alquanto diversa da quella proposta da Quine, perché non richiede un impegno verso il comportamentismo (ora anche nella versione interna, connessionistica) e al contempo non sottoscrive l'idea secondo cui le proprietà mentali sono proprietà cerebrali, mentre il carattere *intrecciato* di questo naturalismo si lasciava esprimere nello slogan: *come la struttura della mente rende possibile l'esperienza della realtà, così la struttura della realtà rende possibile l'esistenza della mente*.

Un simile slogan, che a prima vista rasenta la banalità, si articola invece in una serie di punti qualificanti che banali non sono, tra i quali i più generali sono (I) la tesi che l'*epoché* fenomenologica può essere soltanto locale, nel rispetto di un vincolo che ho chiamato "Prin-

[25] L'espressione "*romance of mathematics*" è usata in Lakoff e Nùñez (2000).

[26] Peruzzi (1994). Parlando di "svolta cognitiva", anch'io, come Lakoff, mi riferivo alla fase dei modelli computazionali guidati dall'idea che i sistemi esperti dell'IA dovessero soltanto essere raffinati per dar conto delle capacità semantiche degli esseri umani.

[27] Questo intento programmatico era esplicitamente dichiarato: "the decisive progress of semantic theory stems from grasping the *depth of surface*, in other words […] the roots of semantics lie in the features of macro-objects described by algebraic and differential topology, and in related dynamical schemes of interaction" (*ibidem*, 357).

cipio dell'Invarianza del Potenziale Referenziale", e (II) il *modello-a-fascio* della mente (ove il concetto di fascio è da intendersi non nel senso di Hume, ma in quello proprio della teoria delle categorie).[28] La congiunzione di (I) e (II) conduce a una forma peculiare di modularità (negata da Lakoff) che, grazie a un'ipotesi più liberale (rispetto alla formulazione di Fodor) circa la taglia dei moduli,[29] si estende al carattere dominio-specifico dei cosiddetti "processi centrali" della cognizione invece di limitarsi ai sistemi di input. Il che è in accordo con un'impostazione dello studio del pensiero già perseguita dai gestaltisti e in disaccordo con l'olismo semantico, nelle sue molteplici versioni (da Quine a Putnam). Il modello-a-fascio si propone infatti di superare la dicotomia fra atomismo e olismo semantico ricorrendo a un'articolata fenomenologia dell'opposizione Locale/Globale che non ne trascura la dialettica con l'opposizione Interno/Esterno (Peruzzi 2002), con immediate ripercussioni sugli argomenti addotti pro o contro il realismo e di conseguenza sugli argomenti pro o contro ciò che Lakoff e Johnson, sulla falsariga di Husserl, intendono come *oggettivismo* (che fa propria una versione particolare del realismo).

4. Spazialità originaria

Una cornice teorica come quella appena tratteggiata fa uso di concetti molto diversi da quelli della semantica formale (da Tarski in poi) cui ampia parte della filosofia del linguaggio fa riferimento e nello stesso tempo mostra che il significato filosofico della teoria delle categorie non è confinato ai fondamenti della matematica[30] ma si misura con l'aderenza alla sorgente delle strutture cognitive, ivi comprese quelle inerenti alla logica. L'idea di fondo della teoria cinestetica del significato è la primarietà dell'*intuizione spaziale,* non più presa come un monolito ma articolata in molteplici componenti interconnessi.

Come lo slogan del naturalismo intrecciato, quest'idea sembra a prima vista non essere una novità. Rispetto ai precedenti appelli all'intuizione (specialmente in ambito fenomenologico) c'è invece una differenza notevole, perché se ne evita un ricorrente difetto: in tali appelli, infatti, il concetto di spazio oscilla fra un senso troppo specifico (postulando la struttura metrica dello spazio euclideo) e un senso troppo generico (postulando un'indeterminata struttura topologica), e, quanto al supporto fisico, l'oscillazione si ripete fra un senso ora sovradeterminato (dalla fisiologia degli organi di senso) ora disincarnato (in base alla multirealizzabilità di uno stesso software).

La gamma di componenti dello spazio fenomenologico è stata documentata nelle indagini empiriche degli psicologi evolutivi e dei percettologi, dotate di un'accuratezza ben maggiore di quella risultante dall'auto-auscultazione rivolta al flusso degli *Erlebnisse*, e ha ricevuto una precisazione formale a partire da Bernhard Riemann, con la nozione di varietà, fino a Alexandre Grothendieck, con la nozione di fascio (e, in seguito, di spazio-senza-punti). È emerso che le componenti possano amalgamarsi in vari modi e soprattutto che la spazialità non è qualcosa di statico, analizzabile con una sorta di grammatica posizionale, bensì ospita variazioni governate da schemi cinestetici. È appunto sullo sfondo di queste acquisizioni che ha preso forma l'idea

[28] (I) e (II) sono introdotti in Peruzzi (1994a). La definizione categoriale di fascio generalizza un concetto più specifico, emerso originariamente in geometria algebrica. Per una formulazione del concetto che ne mette in evidenza l'impatto sulla semantica modellistica, Peruzzi (1991).

[29] Come già proposto da Jackendoff (1987). Si veda anche Sperber (1994).

[30] In tale direzione sono da considerare pietre miliari l'impostazione dialettica di Bill Lawvere, come essa è articolata in Lawvere (1969, 1975) e quella strutturale-fenomenologica accennata in Mac Lane (1986). Ho esposto i motivi di questo significato filosofico della teoria delle categorie nella relazione tenuta al convegno "Ramifications of Category Theory", svoltosi a Firenze nel novembre 2003 (cfr. Peruzzi 2006).

secondo la quale tutto ciò che il linguaggio esprime non sia altro che una sovrastruttura della combinatoria degli schemi. Quest'idea prospetta una *Anschauliche Topologie* di cui la teoria cinestetica del significato è il conseguente sviluppo.[31]

Fin dall'origine, negli anni Quaranta del Novecento, le costruzioni categoriali hanno avuto a che fare con il collegamento sistematico fra algebra e topologia. Anche se in seguito, grazie principalmente ai contributi di Bill Lawvere, la ricerca in teoria delle categorie si è arricchita della dimensione logica, rimaneva da analizzare il *fondamento reale* dei concetti usati in algebra, in topologia e nel relativo collegamento, ove il *fondamento reale* è inteso come *radicamento* corporeo in pattern d'interazione motoria, di propriocezione e d'azione entro lo spazio prossemico. Basta una minima familiarità con nozioni come compattezza, separabilità, connessione, omotopia e omologia per rendersi conto che non si può continuare a trattare l'intuizione (spaziale) come qualcosa di pre-compreso unitariamente. Husserl non ignorava la teoria delle varietà, ma che cosa può esserne di una fenomenologia che ignora un secolo di topologia algebrica e differenziale?

È proprio focalizzando le spie linguistiche del radicamento del significato nella spazialità originaria che ha preso avvio la grammatica cognitiva. Finora, le numerose indagini linguistiche condotte in tale ambito non hanno però sfruttato le potenzialità del linguaggio categoriale e ci si è affidati ad abbozzi teorici privi del necessario rigore. La riformulazione categoriale della semantica permette invece di dare una veste rigorosa alla descrizione degli schemi e dei loro trasferimenti metaforici. Ciò che ne risulta non è una teoria del significato come l'hanno intesa i filosofi analitici, perché è una teoria della genesi del significato a partire da una base[32] governata da princìpi che sono strettamente collegati con i fondamenti della matematica.

La spazialità (nelle sue distinte articolazioni) quale dimensione "costitutiva" primaria dell'intreccio naturalistico e il formato categoriale dei princìpi sono in linea con la filosofia dell'*embodied mind* entro la quale si iscrive la mappatura delle *metafore concettuali* (Lakoff), così come con la *grammatica dei casi* iniziata da Charles Fillmore, la *morfo-genesi semiotica* nel senso di Jean Petitot e, più specificamente, con la *dinamica della forza* nel senso di Leonard Talmy.[33] Tutte queste ricerche illustrano bene che cosa sia una fenomenologia naturalizzata in cui le strutture "geometriche" del linguaggio occupano un posto centrale. Sebbene, come già accennato, tra queste ricerche permangano differenze non trascurabili nella terminologia, nella metodologia, nella rilevanza data ad alcuni fenomeni semantici piuttosto che ad altri, e pure nell'orientamento filosofico, il nocciolo comune risiede nel fatto che tanto l'accesso al significato quanto l'accesso alla pura sintassi sono intesi come *carichi-di-spazialità.* A mio giudizio, questo nocciolo comune può essere precisato affermando che le azioni cognitivamente *salienti* sono quelle che alterano la topologia dello scenario di volta in volta espresso linguisticamente.

A questo punto diventa chiaro il motivo per cui la semantica modellistica e la semantica computazionale vanno incontro alla fallacia simbolica: se (a) non si tiene conto che le strutture della corporeità si organizzano attraverso interazioni spaziali cui le risorse motorie sono essenziali, e (b) si trascura che la costituzione fenomenologica del riferimento si realizza *solo* grazie a tali

[31] Questo modo di intendere la spazialità ha un chiaro debito verso i lavori di René Thom. Il termine *Anschauliche Topologie* mi fu suggerito da John Mayberry con riferimento alla hilbertiana *Anschauliche Geometrie.* Onde evitare equivoci a proposito del termine, occorre tener presente la pervasività del movimento nella caratterizzazione di *pattern* di Posizione e di Forma.

[32] Qui, la nozione di "base" si ispira al significato che il termine ha in geometria algebrica, ma elabora tale significato alla luce della nozione di "specie di livello basico" introdotta da Eleanor Rosch e descritta in Rosch *et al.* (1976).

[33] Per un'esposizione delle idee di Talmy e Fillmore si vedano rispettivamente i riferimenti già forniti. Quanto alla morfo-genesi semantica e alla naturalizzazione della fenomenologia, si vedano rispettivamente Petitot (1985) e Petitot (2000).

strutture, allora (c) si perde ogni traccia delle stesse risorse sfruttate sul piano teorico.[34] Ogni schema basico del riferimento è una *gestalt* cinestetica che invece ne reca traccia e l'insieme di tali schemi permette di evitare la fallacia simbolica.

Tuttavia, in relazione a (c) conviene dire subito che il passaggio da *pattern* spaziali alla forma logica è tutt'altro che semplice. Se per esprimere tale passaggio è stata di scarso aiuto la fondazione insiemistica della topologia, di ancor più scarso aiuto è stato l'atteggiamento scolastico diffuso fra i fenomenologi, preoccupati principalmente di decifrare ed eventualmente sviluppare le indicazioni husserliane su interi e parti nella linea della mereologia, ove è generalmente assente ogni preoccupazione di *costruttività* logica, si elaborano formalismi artificiosi, dotati di minima rilevanza matematica, e si trascurano le trasformazioni *continue* (e relativi invarianti), che sono il cuore della topologia nonché, sul piano empirico, uno strumento indispensabile per l'analisi *qualitativa* dei movimenti (effettivi o immaginati) dei corpi – e del *proprio corpo* in particolare. Più precisamente, ai fini di un'analisi qualitativa le nozioni fenomenologicamente primarie sono quelle che riguardano l'identificazione delle *componenti connesse* di uno spazio, la molteplicità delle *figure* all'interno di ciascuna componente, la relativizzazione dei *punti* ai tipi-di-figure e di trasformazioni, la struttura omotopica dei *cammini*, e infine i tipi-base di *azioni* semanticamente rilevanti, eseguite sugli oggetti immersi nello spazio ambiente e organizzate in gruppi (o monoidi) di operatori.

Ma allora a fornire il quadro concettuale necessario non è la topologia generale, e tanto meno la mereologia, bensì la topologia algebrica; ed è solo a partire da questa che s'imposta una seria analisi fenomenologica dei PATTERN-DI-OGGETTI, dei modi in cui emergono oggetti coesivi e di quali invarianti, rispetto a ben definiti tipi di azioni, siano pertinenti alla stabilità del riferimento di sintagmi nominali e verbali.[35] L'interpretazione degli enunciati di un linguaggio naturale ne risulta fortemente *vincolata* e in parallelo con essa è *vincolata* la categorizzazione. Così facendo, in primo luogo, si supera l'ostacolo costituito dall'altra faccia della fallacia simbolica, cioè, dall'idea di una segmentazione arbitraria di dati grezzi e poi della loro libera raccolta in una gerarchia di tipi indipendente da ogni vincolo percettivo: infatti nomi propri e nomi comuni (per quelli "di massa" la questione è più complessa) si riferiscono sistematicamente a oggetti *stabili* e ciò esige che siano esplicitate le relative condizioni d'invarianza; in secondo luogo, si evita la deriva del movimento fenomenologico, ora verso l'analitica esistenziale e l'ermeneutica, ora verso una filologia autoreferenziale, e ora verso una saldatura con la filosofia analitica che ci riporta daccapo alla fallacia simbolica.[36]

Il punto decisivo della teoria cinestetica del significato è che accanto a *PATTERN*-DI-OGGETTI ci sono *PATTERN*-DI-AZIONI, che coinvolgono mappe (morfismi) da una configurazione (di oggetti) a una successiva nel tempo, conservando almeno una parte della struttura in misura sufficiente da consentire la reidentificazione di ciò cui ci riferiamo. Oggetti e morfismi, presi insieme e opportunamente correlati, definiscono categorie di spazi con azioni e fra tali categorie si possono allora considerare mappe (funtori) che conservano, in vario grado, la struttura da

[34] In relazione ad (a) si veda Hopkins (2001); in relazione a (b), Jackendoff (1987) e Spelke (1990); in relazione a (c), Peruzzi (1994).

[35] Naturalmente, la topologia algebrica ha poi bisogno di essere integrata con gli aspetti differenziali per dar conto della semantica degli eventi.

[36] Non tutti coloro che hanno affrontato in un'ottica fenomenologica il tema dei rapporti fra geometria e logica hanno ignorato la teoria delle varietà e la topologia. Basti pensare a Hermann Weyl. Sulla fecondità degli spunti in Weyl (1968) hanno richiamato l'attenzione John Bell e Giuseppe Longo (vedi nota 11). Il rapporto tra questa linea e quella proposta da Gian-Carlo Rota (cfr. Rota 1997), meriterebbe un discorso a parte. Alcune indicazioni al riguardo si trovano in Tragesser (2000).

una categoria all'altra. La teoria delle categorie fornisce il linguaggio per precisare tutto questo, e non è poco, ma il suo linguaggio non basterebbe a renderla essenziale: quel che fa la differenza rispetto a ogni precedente formalizzazione della semantica è l'esistenza di *costruzioni universali* (in termini di aggiunzioni tra funtori), che permettono di raccordare in maniera uniforme gli elementi (I) e (II) indicati nel § 2.

È stato grazie alla scoperta dell'ubiquità di simili costruzioni in matematica che è stato possibile individuare, in rapporto ai fondamenti della matematica, ciò che contraddistingue la prospettiva categoriale, la cui novità non si riduce alla possibilità di riformulare categorialmente la logica e all'ampliamento dell'universo degli insiemi.[37] Come ho accennato all'inizio, quel che mancava era ancora l'anello di congiunzione tra indagine fondazionale e scienze cognitive. L'anello è fornito da una semantica categoriale *fibrata* su un dominio di *oggetti-e-azioni*, i cui tipi basici (schemi) ospitano le figure salienti sotto il profilo cinestetico.

Nel 1995 ho cominciato a elaborare una simile semantica[38] relativamente a enunciati che l'analisi logica tradizionale considera "atomici", ove i *pattern* intuitivi che definiscono la *base* non si limitano a forme statiche, ma includono forme d'azione intrinsecamente connesse alla corporeità. Dunque, l'*Anschauung* non ha niente a che fare con essenze platoniche; ha, bensì, a che fare con le modalità di categorizzazione dello spazio perisomatico e delle azioni in esso effettuabili che hanno rilievo semantico. Sono però necessari due chiarimenti al quadro teorico impostato nel 1995.

Il primo chiarimento riguarda una possibile obiezione secondo la quale la restrizione a enunciati "atomici" decreterebbe l'inadeguatezza del quadro. La risposta a quest'obiezione (che in effetti mi sarei aspettato, ma non è mai stata sollevata) mi permette di evidenziare i due punti qualificanti dell'intero progetto: a) connettivi e quantificatori sono definiti categorialmente (per aggiunzione) e la predicazione negli enunciati "atomici" è descritta categorialmente come composizione di morfismi; b) la residua struttura logica è nascosta proprio negli enunciati "atomici". In vista di a) e b) il quadro teorico risulta alquanto diverso da quello della semantica modellistica "classica" (tarskiana) o relazionale, a-mondi-possibili, adottata da Montague in poi per gestire il raccordo fra semantica e pragmatica formale.

Il secondo chiarimento riguarda un'altra obiezione (che mi è stata mossa) secondo la quale un quadro del genere *dimentica* tutta la mole di studi compiuti nell'ambito della logica e della filosofia analitica. A questa obiezione rispondo che il quadro teorico qui accennato non è alternativo in tutto e per tutto a quello della semantica tarskiana e alla semantica a-mondi-possibili. L'usuale semantica modellistica è recuperata come caso limite, corrispondente a una logica di entità puntiformi, discrète (su uno sfondo globalmente epochizzato), dall'identità decidibile, in assenza di cambiamenti e relative azioni o in presenza di cambiamenti *in abstracto*, privi di vincoli empirici; pure la semantica computazionale è recuperata come un caso particolare, quale diretta trasposizione metaforica di un ristretto sottoinsieme di schemi pre-posizionali presenti nella semantica di base.

Dunque, anche se la maggiore ampiezza del quadro teorico è necessaria per una coerente integrazione con le indagini semantiche condotte nella grammatica cognitiva, l'impostazione categoriale non espunge l'analisi logica, la teoria classica dei modelli o la semantica procedurale: le considera, piuttosto, tutte e tre come semplificazioni estreme che *dimenticano* i trasferimenti

[37] Già in Peruzzi (1989) era presente un riferimento esplicito ai lavori di Alexandre Grothendieck in geometria algebrica.

[38] Nell'esposizione fattane al decimo congresso *Logic, Methodology and Philosophy of Science* svoltosi nell'agosto del 1995 ed esclusa dai relativi atti, ma in seguito pubblicata come Peruzzi (2000).

metaforici degli schemi inerenti alla base e impediscono quindi di esplicitare tali trasferimenti come processi funtoriali mediante i quali il conferimento-di-senso (l'husserliana *Sinngebung)* si propaga nel linguaggio. La semantica cinestetica intende invece esplicitarli e a tale scopo si dispiega in una costruzione a due strati, indicati come la Base e il Lifting. La Base è una strutturazione dell'ambiente fisico, focalizzata su entità percettivamente salienti (specie basiche di oggetti, qualità, movimenti e azioni – e dunque, in particolare, schemi) con relativi puntatori (localizzatori polimorfi, come, per esempio, gli indicali). Il Lifting consiste in un'ampia classe di sollevamenti (culturalmente selezionati) dei componenti costitutivi dalla Base ai più diversi dominî cognitivi.[39]

La teoria delle categorie è stata ironicamente definita un *abstract nonsense.* Può esser definita così anche la teoria cinestetica del significato? In entrambi i casi, il punto è che l'incremento nell'astrazione finisce per ripagare. Nel caso della teoria cinestetica, l'incremento richiesto per descrivere la costruzione a due strati garantisce la flessibilità necessaria per aderire alla varietà dei modelli metaforici. Il raccordo fra la cornice matematica e il vasto repertorio di *case studies* che è stato raccolto a partire dagli anni Ottanta sta nella nozione di *schema*, in quanto identifica, in un'ipotetica scala dal Concreto all'Astratto, il livello di massima generalità che preserva l'intuibilità immaginativa.

Per esempio, lo schema del CAMMINO (DA-A) lascia indeterminata la traiettoria, la velocità, le modalità motorie e il tipo di mezzo fisico (si pensi a verbi come *andare* e *attraversare*) eppure, una volta congiunto con lo schema del CONTENITORE (DENTRO/FUORI), lo schema del CAMMINO consente di distinguere, senza la mediazione di passaggi inferenziali, uno stato motorio da un altro (*entrare in* / *uscire da* / *restare a*) attraverso corrispondenti configurazioni nell'immaginazione. Il ristretto numero di costituenti *tematici* (i ruoli attanziali per la tradizione francese) in ogni schema (SOGGETTO-FIGURA che passa da una posizione SORGENTE a una posizione BERSAGLIO, OPERATORE che AGISCE su un OGGETTO-FIGURA con uno STRUMENTO, ecc.), unitamente al ristretto numero degli schemi-base, massimizza l'efficienza del trasferimento di significato da un dominio all'altro, perché sono sempre gli stessi *pattern* che plasmano i più vari universi di discorso.

In parallelo, cambiano i termini della contrapposizione fra contesti detti "estensionali" e contesti detti "intensionali", così come si è precisata nella filosofia analitica e nella riformulazione che il concetto di intenzionalità proposto da Brentano vi ha trovato. Ma qui non intendo entrare nel merito perché il discorso esigerebbe quei dettagli tecnici che finora ho cercato di minimizzare. Mi limito a osservare che l'impostazione categoriale non conserva la forma usuale del principio di estensionalità, mentre consente di distinguere diverse forme del principio di composizionalità (Peruzzi 2005); in conseguenza di ciò, alcuni caratteri dell'analisi logica cambiano. Del resto, questo cambiamento era preannunciato nella risposta alla prima delle due obiezioni su menzionate: *se* le proposizioni complesse sono dovute all'impiego di connettivi e quantificatori[40] i quali risultano definibili in termini categoriali mediante aggiunzioni, *se* le proposizioni che per molto tempo sono state indicate come "atomiche" risultano (proprio come gli atomi fisici) strutture più complesse di quanto l'usuale analisi logica lascia pensare, e *se* in tali proposizioni la predicazione risulta dal Lifting o dalla combinatoria degli schemi, *allora* ci sono strutture formali del pensiero che non trovano riscontro nei manuali di logica.

[39] Perciò il quadro risultante esige una revisione profonda delle idee di Husserl sulla composizione dei noemi. Gli aspetti del Lifting da strutture su un dato spazio-base a strutture su spazi fibrati su di esso, sono spiegati in Peruzzi (2004b).

[40] Gli operatori modali ed epistemici richiedono un'analisi categoriale che solo in parte è stata compiuta.

Quanto alle condizioni di verità di qualsiasi proposizione (asserita), esse restano relative al dominio di volta in volta inteso, ma l'unica finestra d'accesso alla *Bedeutung* è data da schemi di oggetti-e-azioni che, come già osservato, corrispondono a una ristretta gamma di *pattern* di interazione corporea nel mondo, o meglio, nel mondo fenomenologicamente "originario", cioè quello mesoscopico degli oggetti-alla-mano. Perciò le proposizioni apparentemente atomiche perché formate con costanti individuali, simboli funzionali primitivi e predicati appunto "atomici" non sono più tali dal punto di vista di una semantica che tiene conto della fibrazione di volta in volta effettuata sulla spazialità originaria (primitiva, schematica) e, onde evitare equivoci, ho accolto un suggerimento di Leonard Talmy e ho parlato di proposizioni *semplici* per indicare quelle atomiche dal punto di vista cinestetico. Una proposizione *semplice* è quella che non fuoriesce dalla Base quanto ad apparato concettuale e che si riferisce a una configurazione caratterizzata esclusivamente dalla COLLOCAZIONE (STATO) di una FIGURA rispetto a uno SFONDO o dal CAMMINO (CAMBIAMENTO-DI-STATO) di una FIGURA rispetto a uno SFONDO (che può anche ridursi a una o più FIGURE) attraverso un'AZIONE elementare di uno o più OPERATORI. Qui, per FIGURA s'intende un oggetto percettivo unitario, lo STATO e il CAMBIAMENTO-DI-STATO sono intesi essere di livello basico e gli OPERATORI sono intesi come agenti che possono servirsi di strumenti, ove agenti e strumenti sono altre FIGURE.

L'integrazione fenomenologica dell'analisi del linguaggio entra dunque in gioco non a cose fatte, dopo una sintassi svincolata dalla semantica e dopo un'astratta semantica formale, ma fin dall'inizio, cioè, nell'analisi tipata della predicazione e nel suo radicamento in proposizioni *semplici* della Base. La traccia linguistica del radicamento si manifesta nella struttura *pre-posizionale* (non necessariamente esplicitata mediante autonome preposizioni) che corrisponde allo schematismo. È così che la spazialità pervade il pensiero.

Finché si escludono dall'analisi logica le risorse di localizzazione tropica (che in lingue come l'italiano sono espresse da preposizioni come IN, A, DA ecc.) e la relativa spazializzazione del tempo (di cui sono spia espressioni come: DA *ieri* A *oggi*, *il passato è* DIETRO *di noi*/*il futuro è* DAVANTI *a noi*), vengono oscurate le condizioni del riconoscimento di stati di appartenenza, possesso, inclusione, che con *nonchalance* sono sfruttate in continuazione nell'analisi logica della sintassi e della semantica. Finché si trascura la radice percettiva della FRONTIERA tra regioni, il soggetto e l'oggetto di una proposizione non possono emergere come figure su uno sfondo. E finché non si tiene conto della distribuzione LOCALE/GLOBALE delle qualità sensoriali e i relativi *pattern* cinestetici, la comprensibilità delle azioni viene meno.

Insomma, la più elementare forma logica è carica di una concreta spazialità (implicita) che si conserva nel pensiero "astratto" e proprio per questo la fallacia simbolica risulta una mezza verità, perché fa supporre che la manipolazione simbolica manchi di qualcosa che le è estraneo, mentre invece tale manipolazione ne è solo l'ombra. L'apparente difficoltà di comprendere una proposizione quanto più è formale è la traccia del ruolo sincategorematico degli schemi-di-base: infatti, dato che essi non sono un componente fra altri della concettualizzazione bensì il mezzo che rende possibile l'unità semantica di qualunque proposizione. Tematizzarli nella loro purezza è come educare un bambino che ha imparato a guardare nella direzione di un dito puntato, a tornare all'inizio, cioè, a guardare il dito – un'educazione che ha aspetti in comune con gli esercizi meditativi zen.

L'ubiquità (!) degli schemi e la loro preservazione nel Lifting mostrano che la concettualizzazione dipende essenzialmente dalla strutturazione dello spazio (e del tempo). Il che, nel lessico kantiano, equivale a dire che l'analitica implica l'estetica, quindi non sussiste autonomamente. Di conseguenza, la fallacia simbolica non è neppure possibile e, per usare un altro termine del lessico kantiano, l'argomento con cui siamo giunti a tale conclusione può esser

definito come "trascendentale", perché risale alle *condizioni di possibilità* della competenza semantica. Nell'architettura a due strati della semantica, l'analisi delle pro-posizioni comincia dunque dalla struttura pre-posizionale ed è così che prende forma la "genealogia" della logica che Husserl intendeva ricostruire,[41] mentre il recupero dell'intenzionalità "originaria" inizia con l'*Anschauliche Topologie* incentrata su schemi d'oggetto-e-azione.

La distinzione di Frege tra *Sinn* e *Bedeutung* e la successiva distinzione fra intensione ed estensione (nella semantica a-mondi-possibili) sono inadeguate a tale recupero perché tanto quella *Art des Gegebenseins* che Frege si limita a nominare allorché descrive il *Sinn,* quanto l'intensione, che da Carnap a Montague è una funzione definita su un insieme arbitrario di mondi possibili, presuppongono l'intero processo di costituzione fenomenologica di ciò che è identificabile come oggetto di riferimento e dimenticano i *pattern incorporati* coinvolti in tale processo. Perciò, l'opposizione che conta non è più quella tra semantica estensionale e semantica intensionale così come l'opposizione si è configurata sullo sfondo della teoria degli insiemi, ma tra una semantica a uno strato (onnicomprensivo) e una a due strati – e solo due perché il Lifting metaforico non si itera –, la prima insiemistica e la seconda categoriale.

Nella semantica a due strati, la Base è fissa[42] mentre le possibili articolazioni del Lifting sono ampiamente variabili. Conseguentemente, la lunga polemica fra intensionalisti ed estensionalisti risulta fuorviante: messo da parte ogni vincolo estensionale, la variabilità del Lifting sarebbe totalmente indeterminata. Un controllo sulla variazione è invece possibile in termini di un'estensionalità che, essendo relativa alle FIGURE elementari di ciascun dato universo di discorso, risulta dominio-specifica, ma ci sono anche concetti polimorfi (nel senso del lambda-calcolo) la cui natura schematica è ancorata alla Base e il cui "riempimento" (*Erfüllung*, nel lessico husserliano) non meno dominio-specifico, varia attraverso i contesti cognitivi che ciascun modello metaforico privilegia.

Riassumendo, la teoria cinestetica prevede che l'architettura del sistema semantico a due strati sia governata da *universali* che vincolano il Lifting dal primo al secondo e inoltre prevede l'esistenza di quattro diversi tipi di universali:

- EIDETICI (schemi, in qualità di attrattori di posizione e d'interazione),
- TEMATICI (*themata* o *ruoli* attanziali, in qualità di dimensioni referenziali primitive),
- FIGURALI (attrattori di forma per oggetti e azioni),
- COMBINATORI (*pattern* di composizione e scomposizione).

In particolare, il Lifting riguarda anche i *pattern* d'inferenza logica, come metaforizzazioni di posizioni e cammini in uno spazio discretizzato, descrivibile come una categoria in cui gli

[41] Seppure con oscillazioni sul senso da dare a tale genealogia. Ci sono, infatti, notevoli differenze prospettiche tra Husserl (1929) e Husserl (1939).

[42] Un'epistemologia generalizzata, cioè, capace di trattare il cambiamento-di-base, dovrebbe preoccuparsi di stabilire se esistono "concetti puri" invarianti rispetto a forme dell'intuizione diversamente strutturate. Facendo l'ipotesi che tali invarianti coincidano con gli schemi, la differenza tra sistemi intelligenti in natura starebbe nella diversa dotazione di schemi che è selezionata da un diverso accoppiamento con l'ambiente.

oggetti sono pro-posizioni e i morfismi sono deduzioni. In questo modo, il dominio logico si costituisce come *categoria* (in senso matematico) di *noemi fondati*.[43]

5. Elementi di confronto

L'idea della metafora come Lifting di struttura a partire da una stessa Base, finita, di schemi è un tratto caratteristico ma non esclusivo della teoria cinestetica. Ciò che la distingue dalle altre linee seguite all'interno della grammatica cognitiva risiede nell'impiego dei concetti categoriali, nell'impegno fondazionale nei confronti della matematica e nel progetto di una geometrizzazione della logica. Questi caratteri comportano non marginali differenze rispetto al quadro che George Lakoff e Rafael Nùñez hanno offerto di alcuni importanti aspetti del linguaggio matematico.

Ai fini di un'equa comparazione con tale quadro, e con altri ancora che si collocano nell'orizzonte di una fenomenologia naturalizzata,[44] sarebbe quindi necessario entrare in dettagli tecnici che preferisco omettere. Mi limito a ricordare che il precursore di un'impostazione geometrica della semantica è stato René Thom, con il suo studio topologico-dinamico della *cattura* (e del suo inverso: l'*espulsione*). In particolare, Thom (1980) è stato anche il primo a impostare il problema di una classificazione della gamma dei tipi di proposizioni definite "semplici", anche se ha trascurato la rilevanza della logica per un'adeguata teoria della semiogenesi: una trascuratezza peraltro diffusa anche nella letteratura fenomenologica e, più recentemente, negli studi di filosofia della mente orientati all'idea di una *embodied mind*.

La prospettiva di Thom è stata ripresa e sviluppata autonomamente da Petitot (2000), il quale ha avanzato l'ipotesi che i *ruoli* posizionali in regioni localizzate dello spazio-tempo agiscano come *archetipi* della struttura del linguaggio, in modo da reinterpretare, prima, la grammatica dei casi e, poi, le idee espresse da Husserl nella terza delle *Logische Untersuchungen* e in *Ideen I* (per quanto concerne la nozione di adombramento percettivo). Petitot ha successivamente proposto una "eidetica" correlata alle strutture neurali e, per alcuni aspetti, la linea seguita da Petitot è compatibile con l'impostazione categoriale della teoria cinestetica. Sebbene impostazioni diverse non necessariamente portino a divergenze nei risultati, in questo caso una differenza significativa è la priorità assegnata nella teoria cinestetica a concetti di topologia algebrica, mentre questa priorità non trova riscontro nella teoria morfogenetica avanzata da Petitot.

D'altra parte, il collegamento con *pattern* neurali la cui peculiarità marca la distanza dalla maggior parte dei modelli computazionali della cognizione, può segnare un vantaggio per tale orientamento, così come per la NTL, rispetto alla teoria cinestetica. Se l'aggancio degli

[43] In maniera inaspettata, quest'impostazione ha rivelato l'opportunità di introdurre un ancor più radicale concetto: quello di *precategoria*, vedi Peruzzi (1994b), corrispondente all'assenza di uniformità nella definizione delle mappe d'identità associate a ciascun oggetto-proposizione, al fine di descrivere più finemente il comune ricorso allo schema del CAMMINO nell'ambito logico, rendendo al tempo stesso la nozione di oggetto non più indipendente dai particolari tipi di processi espressi come morfismi. Dietro a questa generalizzazione si può scorgere la ripresa dell'idea di Eulero del mondo come fluido, in cui le entità sono concrezioni la cui stabilità è relativa a specifiche trasformazioni, solo che ora si tratta del mondo-così-come-appreso-originariamente (per usare il lessico fenomenologico). Ciò significa che in una precategoria viene meno l'uniformità delle mappe di identità, in quanto elementi neutri per la composizione, benché relativizzati a ciascun oggetto. Infatti, in una categoria, per ogni oggetto A c'è una mappa u da A ad A tale che per ogni morfismo f di dominio A, vale che $fu = f$, e tale mappa è appunto quella che si chiama "identità-su-A"; mentre in una precategoria abbiamo soltanto che per ogni oggetto A e per ogni simile f esiste un u da A ad A tale che $fu = f$.

[44] Si pensi innanzitutto alla *morfologia* della percezione visiva, Petitot (1979), e alla ricerche di biologia matematica, come per esempio quelle documentate in Bailly e Longo (2004), nonché, in ambito filosofico, all'analisi degli schemi "estetici" condotta da Fabrizio Desideri (cfr. Desideri 2018).

aspetti geometrici all'emergere di regimi stabili in un sistema dinamico complesso è un punto a favore della linea morfogenetica e di quella centrata sul Lifting categoriale, un vantaggio della grammatica cognitiva, così come presentata dai linguisti americani, sta nell'accurata e vasta documentazione di un'ampia gamma di *pattern* nelle più diverse lingue naturali. Anche in Europa, del resto, c'è stato in tal senso un non meno importante contributo con il progetto IMAGACT, coordinato da Massimo Moneglia (2013). Benché simili linee di ricerca in linguistica, così come quelle nelle neuroscienze cognitive siano in continua evoluzione e quindi ogni giudizio comparativo possa essere soltanto provvisorio, l'impostazione categoriale della teoria cinestetica si configura come una stabile risorsa ai fini di un'integrazione delle evidenze raccolte.

Ormai dovremmo aver capito che per fare dei passi avanti nelle indagini di semantica cognitiva non è più possibile fermarsi a chiosare gli spunti dei padri fondatori o richiamarsi ai paradigmi logici e linguistici sui quali per ampia parte del Novecento si è imperniata la discussione teorica e ai quali si è attenuta gran parte della filosofia del linguaggio. Si tratta piuttosto di operare, all'interno della semantica cognitiva, una selezione dei modelli teorici secondo i classici criteri di coerenza, adeguatezza descrittiva e potere esplicativo. Un'attendibile selezione richiede un più assiduo confronto di idee, rispetto all'attuale scenario, fra chi fa ricerca in ambiti attigui, ma questo confronto richiede una rigorosa precisazione dei modelli teorici della semantica. Se occorre una più stretta collaborazione fra ricercatori di discipline diverse, che affrontano problemi reciprocamente connessi, affidandosi a opzioni teoriche diverse, occorre anche una cornice formale entro la quale effettuare il confronto e, sotto questo profilo, la cornice qui delineata presenta vantaggi non indifferenti rispetto ad altre cornici.

References

Bailey, David R. 1997. *When Push Comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs.* PhD Thesis. Computer Science Division. Berkeley, CA: University of California.

Bailly, Francis, and Giuseppe Longo. 2004. "Space, Time and Cognition: From the Standpoint of Mathematics and Natural Science." In *Mind and Causality*, ed. by Alberto Peruzzi, 149-198. Amsterdam: John Benjamins.

Bell, John. 2004. "Hermann Weyl's Later Philosophical Views: His Divergence from Husserl." In *Husserl and the Sciences*, ed. by Richard Feist, 173-186. Ottawa: University of Ottawa Press.

Carpenter, Bob. 1998. *Type-logical Semantics.* Cambridge, MA: MIT Press.

Desideri, Fabrizio. 2018. *Origine dell'Estetico. Dalle Emozioni al Giudizio.* Milano: Carocci.

Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge: Cambridge UP.

Fauconnier, Gilles, and Mark Turner. 1998. "Conceptual Integration Networks." *Cognitive Science* 22 (2): 133-187.

Feldman, Jerome, and Srinivas Narayanan. 2004. "Embodied Meaning in a Neural Theory of Language." *Brain and Language* 89: 385-392.

Fillmore, Charles. 1968. "The Case for Case." In *Universals in Linguistic Theory,* ed. by Emmon Bach, and Robert T. Harms, 1-88. New York, NY: Rinehart and Wilson.

Fillmore, Charles. 1982. "Frame Semantics." In *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea, 111-137. Seul: Hanshin Publishing Co.

Fillmore, Charles. 1985. "Frames and the Semantics of Understanding." *Quaderni di Semantica* 6 (2): 222-253.

Freeman, Walter, and Rafael Nùñez. 1999. "Restoring to Cognition the Forgotten Primacy of Action, Intention and Emotion." *Journal of Consciousness Studies*, 6 (11-12): 9-19.

Gallese, Vittorio. 2003. "A Neuroscientific Grasp of Concepts: From Control to Representation." *Philosophical Transactions of the Royal Society* 358: 1231-1240.

Goguen, Joseph. 1999. "An Introduction to Algebraic Semiotics, with Application to User Interface Design." In *Computation for Metaphor, Analogy, and Agents*, ed. by Chrystopher Nehaniv, 242-291. Berlino: Springer.
Gruber, Jeffrey. 1976. *Lexical Structures in Syntax and Semantics.* Amsterdam: North-Holland**.**
Hopkins, Brian. 2001. "Understanding Motor Development: Insights from Dynamical Systems Perspectives." In *Handbook on Brain and Behaviour in Human Development*, ed. by Alex F. Kalverboer, and Albert Gramsbergen, 591-620. Dordrecht: Kluwer.
Husserl, Edmund. 1929. *Formale und transzendentale Logik.* Halle: Max Niemeyer.
Husserl, Edmund. 1939. *Erfahrung und Urteil: Untersuchungen zur Genealogie der Logik,* ed. by Ludwig Landgrebe. Praga: Academia.
Jackendoff, Ray. 1983. *Semantics and Cognition.* Cambridge, MA: MIT Press.
Jackendoff, Ray. 1987. *Consciousness and Computational Mind.* Cambridge. MA: MIT Press.
Johnson, Mark. 1987. *The Body in the Mind*. Chicago, IL: The University of Chicago Press.
Kelso, J. A. Scott. 1995. *Dynamic Patterns: The Self-organization of Brain and Behavior*. Cambridge, MA: MIT Press.
Lakoff, George. 1987. *Women, Fire and Dangerous Things.* Chicago, IL: University of Chicago Press.
Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By.* Chicago, IL: University of Chicago Press.
Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh.* New York, NY: Basic Books.
Lakoff, George, and Rafael Nùñez. 2000. *Where Mathematics Comes From.* New York, NY: Basic Books.
Langacker, Ronald W. 1987/1991. *Foundations of Cognitive Grammar.* vol. I-II. Palo Alto, CA: Stanford UP.
Lawvere, Francis W. 1969. "Adjointness in Foundations." *Dialectica* 23: 281-296.
Lawvere, Francis W. 1975. "Continuously Variable Sets; Algebraic Geometry = Geometric Logic." In *Logic Colloquium '73*, ed. by Harvey E. Rose, and John C. Shepherdson, 135-156. Amsterdam: North-Holland.
Lawvere, F. William, and Steve Schanuel. 1997. *Conceptual Mathematics.* Cambridge: Cambridge UP.
Longo, Giuseppe. 2001. "Space and Time in the Foundations of Mathematics, or Some Challenges in the Interactions with Other Sciences." <http://www.di.ens.fr/users/longo> (06/2021).
Mac Lane, Saunders. 1986, *Mathematics. Form and Function.* Berlin: Springer.
Macnamara John, and Gonzalo Reyes. (eds). 1994. *The Logical Foundations of Cognition.* Oxford: Oxford UP.
Mancin, Marco. 2020. *Organismo e Significato: Metodi Topologico-categoriali nelle Neuroscienze.* Tesi di laurea magistrale in Scienze Filosofiche. Università di Firenze.
Moneglia, Massimo. 2011. "IMAGACT. A Multilingual Ontology of Action based on Visual Representations." In *Human Language Technologies as a Challenge for Computer Science and Linguistics*, ed. by Zigmunt Vetulani, Hans Uztoreit, and Marek Kubis, 592-593. Poznan: Fondacja Uniwersytetu im A. Mickiewicza.
Peruzzi, Alberto. 1981. "Logica, Semantica, Fenomenologia." *Antologia Vieusseux* 61-62: 71-82.
Peruzzi, Alberto. 1987. "La Teoria delle Descrizioni: Questo Paradigma della Filosofia." *Epistemologia* 10: 205-226.
Peruzzi, Alberto. 1988. *Noema.* Milano: Franco Angeli.
Peruzzi, Alberto. 1989. "Towards a Real Phenomenology of Logic." *Husserl Studies* 6: 1-24.
Peruzzi, Alberto. 1991. "Categories and Logic." In *Problemi Fondazionali nella Teoria del Significato*, ed. by Gabriele Usberti, 137-211. Firenze: Olschki.
Peruzzi, Alberto. 1994. "Constraints on Universals." In *Philosophy and the Cognitive Sciences*, ed. by Roberto Casati, and Barry Smith, 357-370. Vienna: Hölder-Pichler-Tempsky.
Peruzzi, Alberto. 1994a. "From Kant to Entwined Naturalism." *Annali del Dipartimento di Filosofia* 9: 225-334.
Peruzzi, Alberto. 1994b. "On the Logical Meaning of Precategories." <https://unifi.academia.edu/AlbertoPeruzzi/Papers> (06/2021).
Peruzzi, Alberto. 1996. "Orme nel Silicio, Orme nella Storia." *Paradigmi* 42: 535-553.

Peruzzi, Alberto. 1996a. "Action of Structures, Structure of Actions." *Axiomathes* 7: 325-348.
Peruzzi, Alberto. 1999. "An Essay on the Notion of Schema." In *Shapes of Form,* ed. by Liliana Albertazzi, 191-243. Amsterdam: Kluwer.
Peruzzi, Alberto. 2000. "The Geometric Roots of Semantics." In *Meaning and Cognition*, ed. by Liliana Albertazzi, 169-201. Amsterdam: John Benjamins.
Peruzzi, Alberto. 2002. "ILGE-interference Patterns in Semantics and Epistemology." *Axiomathes* 13: 39-64.
Peruzzi, Alberto (ed). 2004. *Mind and Causality.* Amsterdam: John Benjamins.
Peruzzi, Alberto. 2004a. "Causality in the Texture of Mind." In *Mind and Causality,* ed. by Alberto Peruzzi, 199-228. Amsterdam: John Benjamins.
Peruzzi, Alberto. 2005. "Compositionality up to Parameters." *Protosociology* 21: 41-66.
Peruzzi, Alberto. 2005a. "Il Lifting Categoriale dalla Topologia alla Logica." *Annali del Dipartimento di Filosofia* 11: 51-78.
Peruzzi, Alberto. 2006. "The Meaning of Category Theory for 21st Century's Philosophy." *Axiomathes* 16: 425-460.
Peruzzi, Alberto. 2017. "Schemi cinestetici e metafore nella matematica." In *Verità, Immagine, Normatività. Truth, Image, and Normativity,* ed. by Pier Luigi Lecis, Giuseppe Lorini, Vinicio Busacchi, Pietro Salis, and Olimpia Loddo, 47-59. Macerata: Quodlibet.
Petitot, Jean. 1979. "Hypothèse localiste et théorie des catastrophes. Note sur le débat." In *Théorie du Langage, Théories de l'Apprentissage*, ed. by Massimo Piattelli Palmarini, 516-524. Parigi: Seuil.
Petitot, Jean. 1985. *Morphogenèse du Sens.* Parigi: Presses Universitaires de France.
Petitot, Jean. 2000. "Morphological Eidetics for Phenomenology of Perception." In *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, ed. by Jean Petitot, Francisco Varela, Jean-Michel Roy, and Bernard Pachoud, 330-371. Stanford, CA: Stanford UP.
Port, Robert. F., and Tim van Gelder (eds). 1995. *Mind as Motion. Explorations in the Dynamics of Cognition*. Cambdrige, MA: MIT Press.
Putnam, Hilary. 1999. *The Threefold Chord: Mind, Body and World.* New York, NY: Columbia UP.
Regier, Terry. 1996. *The Human Semantic Potential. Spatial Language and Constrained Connectionism.* Cambridge, MA: MIT Press.
Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Braem. 1976. "Basic Objects in Natural Categories." *Cognitive Psychology* 8: 382-439.
Rota, Gian-Carlo. 1997. *Indiscrete Thoughts.* Boston, MA: Birkhäuser.
Spelke, Elizabeth. 1990. "Principles of Object Perception." *Cognitive Science* 14: 29-56.
Smith, Barry. 1992. "Le strutture del Mondo del Senso Comune." *Iride* 9: 22-44.
Sperber, Dan. 1994. "The Modularity of Thought and the Epidemiology of Representations." In *Mapping the Mind: Domain Specificity in Cognition and Culture*, ed. by Lawrence A. Hirschfeld, and Susan A. Gelman, 39-67. New York, NY: Cambridge UP.
Talmy, Leonard. 2000. *Toward a Cognitive Semantics.* Cambridge, MA: MIT Press.
Thom, René. 1980. *Modèles Mathématiques de la Morphogenèse.* Parigi: C. Bourgois.
Tragesser, Robert. 2000. "Gian-Carlo Rota and the Phenomenological Philosophy of Mathematics. In Memoriam." *Philosophia Mathematica* 8: 3-8.
Varela, Francisco, Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind. Cognitive Science and Human Experience.* Cambridge, MA: MIT Press.
Weyl, Hermann. 1968. *Gesammelte Abhandlungen*, I-IV. Berlino: Springer.
Winograd, Terry, and Fernando Flores. 1986. *Understanding Computers and Cognition.* Boston, MA: Addison Wesley.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: R. Vergari (2021) From polysemy to semantic change: remarks on the lexeme 'miṣwà' in Biblical and Mishnaic Hebrew. *Qulso* 7: pp. 233-262. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12010

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

From polysemy to semantic change: remarks on the lexeme *miṣwà* in Biblical and Mishnaic Hebrew*

Romina Vergari
University of Florence (<romina.vergari@unifi.it>)

Abstract:

The paper discusses the semantic shifts of the lexeme *miṣwà* from Biblical Hebrew to Mishnaic Hebrew. This lexeme encodes some crucial notions of the Hebrew and Jewish value system, as those of commandment and duty. The linguistic data presented will offer a clear example of how the study of the late stratum of the biblical language can shed light on the diachronic semantic shift of the noun. Namely, Late Biblical Hebrew displays some significant semantic innovations, in the form of synchronic contextual variants, that will be fully developed in the post-biblical strata of the Hebrew language to the point of becoming the core meaning of the lexeme.

Keywords: *Diachrony, Mishnaic Hebrew, Polysemy, Semantics of Ancient Hebrew*

Introduction

The present paper is intended to offer a contribution to the semantic study of the lexeme *miṣwà*, which encodes one of the most central notions of the Hebrew and Jewish culture. The linguistic data presented stem from a systematic corpus-based distributional analysis conducted within the Historical-narrative Biblical Hebrew and within the Mishnaic Hebrew.[1] In particular,

* It is my pleasant duty to thank Dr. Hallel Baitner for the inspiring remarks on a first draft of the present paper presented at the Oxford Seminar in Advanced Jewish Studies held in the first half of 2018.

[1] The Mishnà is a collection of legal opinions which became the foundation document of rabbinic Judaism. Compiled in 200 C.E. in Eretz Israel by the patriarch Judah haNasi and his school, the Mishnà comprises the legal statements of the tannaim, i.e. rabbis, and the sages they considered to be their forebears, from Hellenistic times to the early 3rd cent. CE. This material, expressed in a spare post-biblical Hebrew, is arranged in 63 tractates divided into six orders: *Zeraim* ('seeds'), dealing with agricultural matters; *Mo'ed* ('set

the survey aims at showing how the later linguistic strata of BH display innovations in the form of peripheral contextual semantic variants, which will become crucial for the diachronic developments of the lexeme.[2] I have focused my research on the Historical-narrative BH[3] as this language variety[4] often bears witness of linguistic innovations coming from non-literary uses of the language, especially compared to the poetic and the cultic-legal functional languages, which are generally more conservative and archaizing.

1. The usage of miṣwà *in Historical-narrative BH*

Before tackling the examination of the sense-nodules[5] activated by the usage of the noun *miṣwà* in historical-narrative BH, it is useful to make a few overall observations on its distribution and frequency and its syntagmatic features. The noun occurs 64 times in SBH1 (21 of them in the singular and 43 in the plural), and 38 in LBH1 (22 of them in the singular and 16 in the plural).[6] If we normalize the corpora of SBH1 and LBH1 per 10,000 words, we can easily observe that the normalized frequency ratio of *miṣwà* increases considerably from SBH1 to LBH1, going from 5.27 to 8.91.[7] This rise, moreover, concerns mainly the singular (from 1.72 to 5.16), while the plural remains substantially stable (ranging from 3.54 in SBH1 to 3.75 in LBH1).

times'), on the observance of festivals; *Našim* ('women'), primarily on relations between women and men; *Nezikin* ('damages'), on civil and criminal law; *Kodašim* ('holy things'), on sacrifices in the Jerusalem Temple; *Tohorot* ('purities'), on the transfer, avoidance, and removal of ritual pollution. The division into tractates was already more or less established by the 3rd cent., but their arrangement within each order varies in different manuscript traditions. Tractate *ʾAḇot* ('Fathers'), a collection of wisdom sayings by a range of rabbis included within the order *Nezikin*, belongs to a different literary genre from the rest of the Mishnà. It includes a few quotations by rabbis of the generation after Judah haNasi, and may have been added to the Mishnà after its initial redaction; for an introduction see Strack and Stemberger 1992: 119-166, and Cohen 2007: 121-143. For the place of MH within the history of Hebrew language, see Sáenz-Badillos 1993: 161-201.

[2] With regard to the debated topic of diachrony in BH, especially in the domain of lexical semantics, I refer to the works of Avi Hurvitz, starting from his Hebrew University doctoral thesis, Hurvitz 1972, and subsequent works on corpus-based approach to the study of the BH lexicon, as Hurvitz 1995.

[3] For the identification of the Ancient Hebrew functional languages, I refer to the following important works: Vivian 1978; Zatelli 1978; Zatelli 1995, and Zatelli 2004.

[4] A *language variety*, or *lect*, is any intra-linguistic cluster of phenomena that we tend to refer to as dialect, sociolect, stylistic varieties; see Geeraerts and Kristiansen 2019: 150.

[5] For the definition of sense-nodules as relatively autonomous units of sense capable of playing an independent role in various semantic processes, see Cruse 2000: 30.

[6] For an analytical presentation of the data relating to the syntagmatic analysis of the noun in BH, see Appendix 5 in Vergari 2021.

[7] Considering that SBH1 and LBH1 are not corpora of the same size, the number of occurrences of a given textual item does not accurately reflect the relative frequency of it in each corpus. In order to compare corpora (or sub-corpora) of different size, we need then to normalize the occurrences of the item based on the respective total number of words, assumed to be 121,409 for SBH1 and 42,628 for LBH1. The raw frequencies of *miṣwà* are then: SBH1 = 64 per 121,409 words; LBH1 = 38 per 42,628 words. To normalize, we want to calculate the frequencies of our lexical item for each corpus per the same number of words. The convention is to calculate per 10,000 words for smaller corpora and per 1,000,000 for larger ones. In our case, we clearly opt for normalizing per 10,000. Calculating a normalized frequency is a straightforward process. The equation can be represented in this way: 64/121,409 is equal to *x*/10,000. We have 64 occurrences of *miṣwà* per 121,409 words in SBH1, which is the same as *x* (our normalized frequency) per 10,000 words. We can solve for *x* with simple cross multiplication: $x(121{,}409) = 64(10{,}000)$; $x = 64(10{,}000)/121{,}409$. Then, we can say that the normalized frequency ratio (per 10,000) of *miṣwà* is equal to 5.2 in SBH1. Generalizing we can find the normalized frequency of a given lexical item (per 10,000) by applying the following function: $F_N = F_O(10^4)/C$, where F_N is the normalized frequency, F_O the observed frequency, and C the corpus size. For the basic tools of lexical statistics, see Baroni 2008.

In terms of diachrony, *miṣwà* is regarded as a later formation compared to other legal words such as *ḥoq* and *mišpaṭ*, which are attested already in ABH. In terms of etymology, *miṣwà* is a transparent word; it is a nominal derivation from the verbal root *ṣwh* 'to command, to order,'[8] with *m-* preformative added to the verbal stem to produce a noun indicating the action to which the verb points (*nomen actionis*), or more frequently to its result (*nomen rei actae*).[9] Based on the distinction between *syntactic derivation* and *lexical derivation*, the noun can be included in the first class. As expected for these types of derivations, the word changes its lexical category from verb to noun, while the eventive meaning of the root is not touched and the noun retains the same valency of the verb.[10] Given its close connection with the root *ṣwh*, the noun embeds the idea of authority, which turns out to be an inherent feature of its meaning.[11] While the other words of the lexical field of 'rules and regulations' very often derive their authoritative reading from their usage in context[12] – mostly via syntagmatic modulation, suffice it here to recall the important role that the verb *ṣwh* plays in the domain of adnominal relative clauses attached to *ḥoq* or *mišpaṭ* – the substantive *miṣwà* refers to the idea of power *per se*, applying both to humans or divine authority.[13]

When divine authority is at stake, two main patterns of usage can be clearly discerned, with a remarkable impact on the reading's modulation. The first syntagmatic pattern is characterized by the usage of the term in the plural, specified by genitives pointing to God and accompanied by joint terms like *ḥuqqim*/*ḥuqqoṭ*, or *mišpaṭim*. This pattern is typical of the formulaic language of the Deuteronomistic discourse tradition. In cognitive terms, the specific function of this text type is to convey the idea that the teaching of Moses is a unified bounded corpus made of discrete statements conceptualized as 'commandments.' Thus, a relation of meronymy can be envisaged between this contextual reading of *miṣwà* (i.e. *miṣwoṭ*) and the term *toṛà* as it is used within Deuteronomy. The second syntagmatic pattern is characterized by the usage of the term in the singular, accompanied by joint terms like *toṛà*, *ḥuqqim/ḥuqqoṭ*, or *mišpaṭim*, additionally combined with the adnominal demonstrative *zoṭ* or the quantifier *koḷ*. As I will show through the following examples, this pattern's frequency increases considerably from SBH1 to LBH1. When the context triggers this particular reading, *miṣwà* turns out to be a referential synonym of *toṛà*, with remarkable ideological implications.[14] In Deuteronomy and Deuteronomistic discourse tradition, *toṛà* and *miṣwà* appear to function as onomasiological alternatives to name the teaching of Moses in its path of formalization and fixation, and *miṣwà* is chosen precisely to place special emphasis on the authoritative aspect of it. In historical-narrative language thus the body of literature considered authoritative can be conceptualized in a unified manner either as a teaching (mainly an oral teaching in SBH1, and a written text to be expounded, explained,

[8] See *HALOT*, 7899: 1) 'to give an order, to command'; 'to command, instruct, order'; 3) 'to send someone (to a place, for a task)'; *BDB*, 8061: 1) 'to lay,' 'to charge upon'; 2 and 3) 'to charge,' 'to command'; 4) 'to commission'; 5) 'to appoint,' 'to ordain'; for more detailed syntagmatic information see also *DCH* 7: 93-102. Jenni includes this stem among the transitive resultative verbs without basic form *qal*; see Jenni 1968: 246-248.

[9] See Joüon and Muraoka 2006: § 88 L, *e*.

[10] In nouns formed via lexical derivation instead, the change of category also affects the meaning, as in the case of *zbḥ* 'to slaughter for sacrifice,' and *mizbęaḥ* 'altar'; see Panevová 2014: 7. Such a distinction has been set by the seminal work Kuryłowicz 1936.

[11] See Levine *TDOT* 8: 506.

[12] Especially through the usage of adnominal relative clauses.

[13] See *BDB* 8063: 1) 'commandment' of men (vz. of kings); 2) 'commandment of God,' in the singular: 'commandment,' 'code of law'; in the plural 'commandments,' of commands of D and later codes; and *HALOT*, 5540: 'commission,' '(individual) commandment,' '(set of all the) commandments,' 'right'; see also *DCH* 5: 446-448, 'command(ment).'

[14] For a definition of referential synonymy, see Grondelaers, Speelman, and Geeraerts, 2007: 994-995.

and interpreted in LBH1) or as a command to be executed. The latter conceptualization is far from being obvious. It is important to point out, as Levine has done, that it is within the hortatory Deuteronomic discourse tradition that the divine will expressed in the body of Scriptures as a unified textual corpus was initially understood and then transmitted as a command. It is worth stressing that this particular interpretation will be maximized in later rabbinic tradition, especially in halakhic discourse. As I will show, the data emerging from the present corpus-based analysis basically agree with the research in the domain of textual criticism in connecting this specific reading with Deuteronomistic redactional activity.

1. Expression of Divine Authority

1.1 The teaching of Moses as 'Commandment'

In Standard historical-narrative BH, especially within Deuteronomy and Deuteronomistic discourse tradition, many examples can be found of a collective reading of *miṣwà*,[15] which parallels in many respects the usage of the noun *ṭorà*. I will focus on three main text types: *kṭl hammiṣwà* (singular definite plus quantifier)[16]; *hammiṣwà hazzṭt* (singular definite plus adnominal demonstrative)[17]; and the pair *hattṭrà wəhammiṣwà*.[18]

All these syntagmatic types are united by two facts. On the one hand, *miṣwà* occurs without those adnominal modifiers (pronominal suffixes or genitives) that are required for encoding the complements of eventive nouns. This fact suggests that the substantive is slowly changing its semantic type. It is formed through a syntactic derivation to indicate the process or the act of commanding and is becoming a referential noun that points to a complex object. On the other hand, the schemes under scrutiny convey a similar interpretation of *miṣwà*, which does not correspond to the uniplex reading 'one single commandment' but rather to an abstract unified notion corresponding to the revelation of the divine will as a whole.

I begin my analysis with the text type *kol hammiṣwà*, in which *kṭl* functions as the universal quantifier 'all,' 'whole' and *hammiṣwà* as its determiner. The combination turns out to be quite peculiar, if one compares the standard usage of *miṣwà* and other terms for rules and regulations in similar phrases.[19]

[15] A certain number of them are listed in Wienfeld's appendix "Deuteronomic phraseology"; Weinfeld 1972: 320-365; *DCH* distinguishes between 'singular used collectively' (Exod 24:12; Num 15:31; Deut 5:31; 6:1; 7:11; 8:1; 11:8; 15:5; 19:9; Josh 22:3; Ps 19:9; 119:96; 2 Chr 14:3; 31:21; Sir 6:37; 10:19; 15:15; 35:18.23; 37:12; 44:20; 45:5; 1QpHab 5:5; 1QS 8:17; 4QDc 1:6; GnzPs 1:10; and singular for 'one particular command' (1 Sam 13:13; 1 Kgs 13:21; Mal 2:14; Job 23:12; 2 Chr 29:25); see *DCH* 5: 446.

[16] See Deut 8:1; 11:8, 22; 27:1; 31:5 (SBH1); compare also Deut 5:31; 15:5; 19:9 (SBH4).

[17] See Deut 11:22; 30:11 (SBH1); compare also Deut 6:25; 15:5; 19:9 (SBH4).

[18] See Exod 24:12; Josh 22:5; in combination with other terms: 2 Kgs 17:34.37 (SBH1); and 2 Chr 14:3; 31:21 (LBH1). It must be said that additional schemes could be added, that convey a unified conceptualization of the Mosaic teaching (or the divine will) as *command*. On the one hand, some occurrences attest the usage of the noun in the singular, combined with other terms for divine precepts in plural, suggesting a semantic relationship of meronymy between them: *ʾt ḥmṣwh wʾt ḥḳym wʾt ḥmšṗṭym* (Deut 7:11). On the other hand, cases in which the term in the singular is specified by a relative clause with the verb *ṣwh* trigger the idea that the divine will be revealed through the mediation of Moses is a command (Deut 27:1). To this conceptualization, must be added the phrase *mṣwt Mšh* (2 Chr 8:13). Finally, the idea that the revelation of the divine will is a command *per se* is definitively sanctioned by expressions as *mṣwt YHWH* (or *ʾlḥym*) (Josh 22:3; 1 Sam 13:13; Ezra 10:3).

[19] See Netzer 2013: 313; see *HALOT*, 4240, namely the meaning listed as seventh: "*kl* preceding collective

Normally, *miṣwà* combines with *kṣl* in the plural, yielding the multiplexing reading 'all the commandments,' as in the following example:[20]

(1) Deut 28:1
whyh ʾm šmwʿ tšmʿ bqwl YHWH ʾlhyk lšmr lʿśwt ʾt kl mṣwtyw ʾšr ʾnky mṣwk hywmwntnk YHWH ʾlhyk ʿlywn ʿl kl gwyy hʾrṣ
'And it shall come to pass, if you shall hearken diligently unto the voice of YHWH your God, to observe to do *all his commandments* which I command you this day that YHWH your God will set you on high above all the nations of the earth.' (NKJV)

The same applies to *ḥṣq*, *ḥuqqà* and *mišpaṭ*.[21] These lexical items, nevertheless, occur as determiners of *kṣl* also in the singular, but compared to *miṣwà*, they do not come to designate the whole teaching of Moses. The reading that most frequently arises in context is rather 'one single (specific) instance as a whole' of the type of statements to which each lexeme refers. The reference of such phrases corresponds to a cohesive unit excerpted from a body of statements alike. This phenomenon is observable in the following context:

(2) Num 9:12
kkl ḥqt hpsḥ yʿśw ʾtw
'*according to the whole regulation of Pesaḥ* they shall do it.'[22]

In this passage, the reference of *ḥuqqat* is further bounded by the genitive *happesaḥ*, triggering the reading 'according the whole regulation of *Pesaḥ*.' It is in fact a special set of rules regarded as a unity and singled out from a multiplex body of discrete statutes regulating other matters.[23] The term *mišpaṭ* displays a similar pattern of usage in two instances pertaining to SBH2 and LBH2; in both the noun is further specified, in one case by a suffix:

'all': *kl hʾdm* 'all men' Gen 7:21 (also Num 12:3; Judg 16:17)"; see also *BDB* 4485: "*kl* followed often by a singular, to be understood collectively, whether with or without the article: e.g. 2 Sam 20:22 *wtbwʾ hʾšh ʾl kl hʿm* 'the woman went (to speak) to all the people."

[20] See also Deut 4:6; 28:1.15.45; 30:8; 1 Kgs 6:12; 2 Kgs 17:16; Jer 35:18 (SBH1); and 1 Chr 28:8; 2 Chr 24:20; Neh 10:30 (LBH1).

[21] Concerning *ḥuqqim*, see *ʾt kl hḥqym hʾlh* 'all these statutes' (Deut 4:6), compare also Lev 10:11; Deut 5:31; 6:24; 11:32 (SBH4); concerning *ḥuqqṭ*, see Num 9:3; Deut 6:2 (SBH1); Lev 19:37; 20:22; Ezek 18:19.21; 43:11[x2]; 44:5 (SBH4); Concerning *mišpaṭim*, see Exod 24:3; 1 Kgs 6:38 (SBH1); compare 2 Sam 22:23 (SBH2); Num 9:3; Lev 19:37; 20:22 (SBH4); and Ps 119:13 (LBH2).

[22] Among modern translations, some opt for a collective reading of the phrase *ḥqt hpsḥ*, see: "when they celebrate the Passover, they must follow *all the regulations*" (NIV); "according to *all the ordinances* of the Passover they shall keep it" (NKJV); others provide a unified reading, see "they shall offer it in strict accord with *the law* of the Passover sacrifice" (NJPS); "they will keep it, following *the entire* Passover *ritual*" (NJB); "according to *all the statute* for the Passover they shall keep it" (RSV); "according to *all the statute* of the Passover they shall observe it" (NASB); "the Passover shall be kept exactly as *the law* prescribes" (NEB).

[23] This usage is most likely attested also for *miṣwà*; the expression *ʾt kl hmṣwh ʾšr ʾnky mṣwh ʾtkm hywm* in Deut 27:1 may refer either to the requirement to erect an altar or the requirement to monumentalize the *ṭorà*; but this passage is highly complex in terms of composition, and may reflects multiple additions of different textual material.

(3) Prov 16:33
bḥyq ywṭl ʾt hgwrl wmYHWH kl mšpṭw
'The lot is cast into the bosom and *all its judgment* comes from YHWH.'[24]

in the second case by a governed Nph:

(4) Ps 119:160
wlʿwlm kl mšpṭ ṣdqk
'*each of your righteous judgments* endures forever.'[25]

In the first example, the espression *kol mišpaṭo* designates a single specific response of the *goral*, the lot cast for the decision of questions, whereas the structure *kol mišpaṭ ṣidqe̲ka* in the second example can be explained assuming the distributive universal reading 'each,' 'every' for the quantifier *kol*.

The examples in which *miṣwà* determines *kol* in the singular, on the other hand, deviates decidedly from the pattern sketched above. Firstly, the phrase does not produce the distributive reading 'each commandment' nor the collective one 'all the commandments.' The examples collected suggest rather a unified interpretation pointing to a mass continuous entity, which is bounded only by the relevant adnominal relative clause 'that I command you today.' The noun's referent is thus as extensive as the speeches that Moses is delivering within the framework of the text of Deuteronomy. Accordingly, the usage of *miṣwà* comes to comprise not only the normative or directive sub-sections of these speeches, viz. the rules governing individual subjects, but also the narrative and hortatory parts of them. Such a usage punctuates the redactional interventions scattered throughout Deuteronomy, framing its structure, and expressing the clear ideology of the editors towards the text *in fieri*. In their estimation, the purport of Moses's speeches collected in Deuteronomy must be viewed alternatively as a teaching (*toràh*) or as a command (*miṣwà*). The relevant examples of this reading are listed below.

The current structure of the second oration of Moses (Deut 4:44-28:68)[26] has been regarded as the outcome of a considerable amount of literary activity pertaining to one redactional stratum of the book.[27] The conceptualization of this whole unit as a *miṣwà* appears to fit very well the agenda of the redactors, as the following passage clearly shows.

(5) Deut 8:1
kl hmṣwh ʾšr ʾnky mṣwk hywm tšmrwn lʿśwt lmʿn tḥywn wrbytm wbʾtm wyrštm ʾt hʾrṣ ʾšr nšbʿ YHWH lʾbtykm
'*the whole commandment* that I command *you* (sg.) today, you (pl.) shall be careful to do, that you (pl.) may live and multiply, and go in and possess the land that YHWH swore to give to your fathers.'[28]

[24] See Fox 2009: 623.

[25] Several modern translations render *kl mšpṭ ṣdqk* in plural (NASB; NIV; NJB; NKJV; RSV; NJPS), suggesting a collective reading; see also Weiser's translation "everyone of thy righteous ordinances endures for ever"; see Weiser 1962: 737.

[26] According to Rofé 2002: 1-4.

[27] Many scholars consider the redactor named D_2 responsible for this redactional activity, namely for the opening of the collection (5:1; 6:9); 2), additional portions of the present introduction to chapters 6-11, which originally belonged to the "*toràh*" (7:1-11; 11:22-25), and the overall current structure of the second oration (5:28; 6:1; 8:1; 11:22; 11:32-12:1; 26:16). According to Rofé the objective of the redactor was "to implement a comprehensive legal code, which would secure the status of law of the land through the sanction of royal backing and replace earlier legal compilations or the existing customary law," see Rofé 2002: 6.

[28] Among modern translations, "all the commandment" (RSV), and "all the Instruction" (NJPS) are in line with the reading I propose, while both "all the commandments" (NASB; NJB) and "every command" (NIV; NKJV) suggest a collective interpretation of *kol hammiṣwà*.

As observed by Weinfeld, the shift in person deixis from singular in the first clause (*ʾăšer ʾanoḵi məṣṣawəḵa*), to plural in the rest of the sentence (*tišmərun… tiḥyun urəḇitem uḇaʾtem wirištem*) is replicated in v. 19.[29] This fact may allude to the framing function of both verses, which indeed forms a kind of *inclusio* for chapter 8.[30]

The same degree of literary elaboration can be envisaged in chapter 11, within which the phrase *kol hammiṣwà* plays a significant role as a redactional mark:

(6) Deut 11:8
wšmrtm ʾt kl hmṣwh ʾšr ʾnky mṣwk hywm lmʿn tḥzqw wbʾtm wyrštm ʾt hʾrṣ ʾšr ʾtm ʿbrym šmh lršth
'you shall therefore keep *the whole commandment*[31] that I command you today, that you may be strong, and go in and take possession of the land that you are going over to possess' (RSV)

(7) Deut 11:22
ky ʾm šmr tšmrwn ʾt kl hmṣwh hzʾt ʾšr ʾnky mṣwh ʾtkm lʿśth lʾhbh ʾt YHWH ʾlhykm llkt bkl drkyw wldbqh bw (23) *whwryš YHWH ʾt kl hgwym hʾlh mlpnykm wyrštm gwym gdlym wʿṣmym mkm*
'for if you will be careful to do *all this commandment*[32] that I command you to do, loving YHWH your God, walking in all his ways, and cleaving to him, then YHWH will drive out all these nations before you, and you will dispossess nations greater and mightier than yourselves.' (RSV)

Deuteronomy 11:22 recapitulates the statement expressed in v. 8, but it changes the arguments in favor of loyalty; whereas at the beginning of the section the keeping of the commandment is motivated by the inheritance of the good land and enjoyment of its produce,[33] the reward consists rather in military success in the final reprise of the theme. In this redactional verse, the phrase *kol hammiṣwà* is further specified by the adnominal demonstrative, with an obvious function of discourse deictic. Such an interpretation is attested also in other contexts:

(8) Deut 30:11
ky hmṣwh hzʾt ʾšr ʾnky mṣwk hywm lʾ nplʾt hwʾ mmk wlʾ rḥqh hwʾ
'for *this commandment* that I command you today is not too hard for you, neither is it far off.' (RSV)

Remarkably, this syntagmatic structuring of *miṣwà* parallels that of *ṭorà*. All the data collected suggest an interpretation of *miṣwà* as a continuous bounded entity designating the entire body of the Mosaic teaching in its process of fixation within the book of Deuteronomy, with particular emphasis on its binding force as a commandment that requires first and foremost

[29] See Deut 8:19 *ʾm škḥ tškḥ ʾt YHWH ʾlhyk … hʿdty bkm hywm ky ʾbd tʾbdwn* "if you (sg.) shall forget YHWH your God, and walk (sg.) after other gods, and serve (sg.) them, and worship them, I forewarn you (pl.) this day that you (pl.) shall surely perish."

[30] See Weinfeld 1991: 441.

[31] Compare "all the commandments" (KJV).

[32] Compare "all these commandments" (KJV), that assumes again a collective reading.

[33] This is a typical motif of the Deuteronomistic discourse tradition, see Weinfeld 1972: 341.

observance and obedience.[34] Moreover, such a usage can be traced back to the Deuteronomistic editorial enterprise, representing a peculiar feature of its discourse tradition. The structuring described so far and the underlying ideology deserve a proper place within the Deuteronomistic phraseology and should be integrated in the list of stylistic devices expressing observance of the law and loyalty to the covenant made by Weinfeld.[35]

In addition to what has been observed so far, it must be said that the close connection between *tǫrà* and *miṣwà* as onomasiological alternatives to name the same referent is not only a characteristic typical of Deuteronomy, but it is also found in texts that cannot be directly related to its tradition. To give a clear example, I will now analyze the pair *tǫrà umiṣwà*.[36] This combination occurs both in isolation and within more complex juxtapositions.[37] Interestingly enough, the two terms always agree in number and determination, which is a typical feature of hendiadys.[38] I begin my overview with an emblematic and famous context:

(9) Exod 24:12
wyʾmr YHWH ʾl mšʿlh ʾly hhrh whyh šm wʾtnh lkʾt lḥt hʾbn whtwrh whmṣwh ʾšr ktbty lhwrtm
'And YHWH said unto Moses: 'Come up to me into the mount and be there; and I will give you the tables of stone, *the law* (lit. *the teaching and the commandment*) which I have written, that you may teach them."[39]

Among commentators, Houtman understands *wəhattǫrà wəhammiṣwà* as a hendiadys and renders it accordingly: '(the tablets of stone) containing the binding rules'[40]; Propp, on the other hand, opts for the more literal rendering: '(the stone tablets), the direction and the command.'[41] According to Propp's view, the first *wə* (*wəhattǫrà*) must be understood as explicative.[42] Although he cautiously argues that "it is unclear whether what YHWH proposes to write in 24:12 is the same or a different text,"[43] nevertheless, his translation implies the appositive function of *wəhattǫrà wəhammiṣwà* with respect to *luḥǫt haʾeḇen* and, thus, he takes it as an identity of reference.[44] Another element, moreover, deserves to be taken into due consideration, namely the relative clause *ʾăšer kaṯaḇti ləhǫrǫtam*, which modifies the noun phrase *wəhattǫrà*

[34] See Levine, *TDOT* 8: 509-510.

[35] See Weinfeld 1972: 332-339.

[36] For the plural usage, see Exod 16:28 *miṣwǫtay wətǫrǫtay*, with the multiplexing reading "commandments and instructions" (SBH1); for the singular usage, see 2 Chr 14:3 *hattǫrà wəhammiṣwà*, and 2 Chr 31:21 *uḇattǫrà uḇammiṣwà* (LBH1).

[37] See 2 Kgs 17:34 *kəḥuqqǫtam ukəmišpaṭam wəḵattǫrà wəḵammiṣwà*; and 2 Kgs 17:37 *wəʾet haḥuqqim wəʾet hammišpaṭim wəhattǫrà wəhammiṣwà*.

[38] See the relevant literature on the topic of hendiadys, in particular: Avishur 1972; Kuntz 2004; Lillas-Schuil 2006.

[39] Among modern translations, many understand *miṣwà* as a collective, and render it accordingly, see "I will give you the stone tablets with the law and the commandments" (NIV; NKJV); "I will give you the stone tablets with the teachings and commandments" (NJPS).

[40] See Houtman 2000: 296.

[41] See Propp 2006: 5.

[42] See *GKC* §155, 1a; it must be pointed out, moreover, that both SP *ʾt lḥt hʾbn htwrh whmṣwh*, and LXX τὰ πυξία τὰ λίθινα τὸν νόμον καὶ τὰς ἐντολάς witness a variant without the conjunction before *hattǫrà*.

[43] See Propp 2006: 298-299.

[44] It is important to observe that, unlike what Propp claims, LXX takes only *wəhammiṣwà* as a collective tantamount to plural and not both terms (see τὸν νόμον καὶ τὰς ἐντολάς; see also Vulg. *legem ac mandata*).

wəhammiṣwà. The noun *miṣwà* is not included among the complements of the verb *yrh* (H/1 hiqṭil) 'to instruct,' 'to teach,' while the noun *ṭorà* (etymologically related to this root) is attested twice in this function, both in the domain of relative clauses,[45] and in the domain of verbal phrases.[46] This fact suggests a secondary juxtaposition of the term *miṣwà*, grounded in a process of conceptual identification.

The usage of *miṣwà* and *ṭorà* as a pair is steadily attested across Standard and Late historical-narrative language:

(10) Josh 22:5
rq šmrw mʾd lʿśwt ʾt hmṣwh wʾt htwrh ʾšr ṣwh ʾtkm mšh ʿbd YHWH lʾhbh ʾt YHWH ʾlhykm wllkt bkl drkyw wlšmr mṣwtyw wldbqh bw wlʿbdw bkl lbbkm wbkl npškm
'Only take diligent heed to put in practice *the law*,[47] which Moses the servant of YHWH commanded you, to love YHWH your God, and to walk in all his ways, and to keep *his commandments*, and to cleave unto him, and to serve him with all your heart and with all your soul'

and it is found up to the later linguistic layers of the biblical corpus:

(11) 2 Chr 31:21
wbkl mʿśh ʾšr hḥl bʿbwdt byt hʾlhym wbtwrh wbmṣwh ldrš lʾlhyw bkl lbbw ʿśh whṣlyḥ
'every work that he undertook in the service of the house of God and in accordance with *the law*,[48] seeking his God, he did with all his heart, and prospered.'

1.2 Commandment

Far more frequent and spread across different discourse traditions is the usage of *miṣwà* pointing to a single specific commandment originating from God. Such a reading arises mostly from the usage of the noun in the plural, which expresses the obvious multiplex discrete conceptualization 'commandments.' It is important to observe that the plural occurrences of the term are normally specified either by pronominal suffixes pointing to God (*miṣwọtay*, *miṣwọteḵa*, *miṣwọtayw*)[49], or by the genitive YHWH (*miṣwọt YHWH*)[50], a fact that marks a clear difference with the use described in the previous paragraph.

Two different aspects of this text type deserve special attention, one being formal, and the other referential. Firstly, considering the consonantal shape of the text, the form *mṣwt YHWH*

[45] See Deut 17:11 (SBH4).

[46] See Deut 33:10 (ABH).

[47] Literally "the teaching and the commandment"; compare "the commandment and the law" (NASB; NIV; NKJV; RSV); "the commandments and the Law" (NJB); "the commandments and the laws" (NEB); "the Instruction and the Teaching" (NJPS).

[48] NET translates like this; several translations, however, read *miṣwà* as a collective, compare "the law and the commands" (NIV); "the law or the commandments" (NJB; RSV).

[49] See Gen 26:5; Exod 16:28; Deut 4:40; 8:11; 11:1; 27:10; 28:15.45; 30:10.16; 1 Kgs 2:3; 3:14; 8:58; 9:6; 11:34.38; 2 Kgs 17:13; 23:3; (SBH1); and Ezra 9:10; 9:14; Neh 1:5.9; 1 Chr 28:7; 29:19; 2 Chr 7:19; 17:4; 34:31; Qoh 12:13 (LBH1).

[50] See Deut 10:13 (SBH1); and Ezra 7:11; Neh 10:30; 1 Chr 28:8; 2 Chr 24:20 (LBH1).

is ambiguous in terms of morphological number, it can be read either *miṣwat YHWH* or *miṣwọt YHWH*. Only context, in particular agreement, can help the reader disambiguate such a reading. Otherwise, we must rely on the Masoretic reading tradition. The second aspect concerns the reference of this expression. Its usage suggests that the meaning of *miṣwà* should be regarded as inherently underspecified with respect to the feature "origin of the command." Assuming its vagueness, the term calls for contextual specifications (genitives, relative clauses, pronominal suffixes), which have the main function of focusing the attention of the recipient on the origin of such a command. In other words, the divine origin of the command is not fully lexicalized in the semantics of *miṣwà* in BH as is the case for the English noun *commandment* compared to *command*.[51] Such feature was instead triggered by operations of sematic composition in context. A selection of examples showing this feature follows:

(12) Deut 4:2
lʾ tspw ʿl hdbr ʾšr ʾnky mṣwh ʾtkm wlʾ tgrʿw mmnw lšmr ʾt mṣwt YHWH ʾlhykm ʾšr ʾnky mṣwh ʾtkm
'You shall not add to the word that I command you, nor take from it, that you may keep *the commandments of YHWH your God* that I command you' (RSV)

(13) Deut 11:13
whyh ʾm šmʿ tšmʿw ʾl mṣwty ʾšr ʾnky mṣwh ʾtkm hywm hywm lʾhbh ʾt YHWH ʾlhykm wlʿbdw bkl lbbkm wbkl npškm
'And if you will obey *my commandments* which I command you this day, to love YHWH your God, and to serve him with all your heart and with all your soul (v.14 he will give the rain for your land in its season, the early rain and the later rain, that you may gather in your grain and your wine and your oil)' (RSV)

(14) Deut 11: 26–27
rʾh ʾnky ntn lpnykm hywm brkh wqllh (27) *ʾt hbrkh ʾšr tšmʿw ʾl mṣwt YHWH ʾlhykm ʾšr ʾnky mṣwh ʾtkm hywm* (28) *whqllh ʾm lʾ tšmʿw ʾl mṣwt YHWH ʾlhykm*
'Behold, I set before you this day a blessing and a curse: (27) the blessing, if you obey *the commandments of YHWH your God*, which I command you this day (28) and the curse, if you do not obey the commandments of YHWH your God (but turn aside from the way which I command you this day, to go after other gods which you have not known).' (RSV)

(15) Deut 28:13
wntnk YHWH lrʾš wlʾ lznb whyyt rq lmʿlh wlʾ thyh lmṭh ky tšmʿ ʾl mṣwt YHWH ʾlhyk ʾšr ʾnky mṣwk hywm lšmr wlʿśwt
'And YHWH will make you the head, and not the tail; and you shall tend upward only, and not downward; if you obey the commandments of YHWH your God, which I command you this day, being careful to do them.' (RSV)

[51] The feature "divine origin" is lexicalized in many modern languages that display semantic variance between a vague term "command," and a specific term "divine command", see Italian *comando* vs. *comandamento*; French *ordre* vs. *commandement*; German *Befehl* vs. *Gebote*; Spanish *orden* vs. *mandamiento*.

(16) 2 Kgs 18:6
wydbq bYHWH lʾ sr mʾḥryw wyšmr mṣwtyw ʾšr ṣwh YHWH ʾt mšh
'For he (king Hezekiah) held fast to YHWH. He did not depart from following him but kept *the commandments* that the Lord commanded Moses.' (RSV)

It is useful to mention that the term in the plural occurs often in combination with the quantifier *kọl* within both SBH1 and LBH1:

(17) Deut 28:15
whyh ʾm lʾ tšmʿ bqwl YHWH ʾlhyk lšmr lʿśwt ʾt kl mṣwtyw wḥqtyw ʾšr ʾnky mṣwk hywm wbʾw ʿlyk kl hqllwt hʾlh whśygwk
'But if you will not obey the voice of YHWH your God or be careful to do *all his commandments* and his statutes that I command you today, then all these curses shall come upon you and overtake you' (RSV)

(18) 1 Chr 28:8
wʿth lʿyny kl yśrʾl qhl YHWH wbʾzny ʾlhynw šmrw wdršw kl mṣwt YHWH ʾlhykm lmʿn tyršw ʾt hʾrṣ ḥṭwbh whnḥltm lbnykm ʾḥrykm ʿd ʿwlm
'Now therefore in the sight of all Israel, the assembly of YHWH, and in the hearing of our God, observe and seek out *all the commandments of YHWH your God*, that you may possess this good land and leave it for an inheritance to your children after you for ever.' (RSV)

It is difficult to underestimate the pivotal role that the verb *ṣwh* (0/2, qiṭṭel) (especially within adnominal relative clauses) played in the conceptualization of the will of God as a command or a bounded set of discrete commandments, especially taking into account the fact that the noun *miṣwà* does not occur in the most ancient cultic and legal texts; its place is normally occupied by other terms such as *dəḇarim* (in the plural, in particular in the phrase *dibrę habbərit*)[52], *bərit*,[53] *tọrà*,[54] and *ḥuqqim* (in the plural)[55], used in isolation or in combination to form chains. These lexemes clearly derive their binding value from the syntagmatic relation with the predicate *ṣwh*

[52] Compare Exod 19:7 *wyśm lpnyhm ʾt kl hdbrym hʾlh ʾšr ṣwhw YHWH* "(Moses) acquainted them with everything that YHWH had commanded him"; Exod 35:1 *ʾlh hdbrym ʾšr ṣwh YHWH lʿśwt ʾtm* "these are the things that YHWH has commanded you to do"; Deut 28:14 *wlʾ tqwr mkl hdbrym ʾšr ʾnky mṣwh ʾtkm hywm* "do not deviate to the right or to the left from any of the things that I command you this day"; and Deut 28:69 *ʾlh dbry hbryt ʾšr ṣwh YHWH ʾt mšh* "these are the terms of the covenant which YHWH commanded Moses" (SBH1); see also Jer 11:8 (SBH2); and Lev 8:36; Deut 6:6; 12:28 (SBH4).

[53] Compare Deut 4:13 *wygd lkm ʾt brytw ʾšr ṣwh ʾtkm lʿśwt ʿśrt hdbrym* "(YHWH) declared to you the covenant that He commanded you to observe, the ten commandments"; Josh 7:11 *wgm ʿbrw ʾt bryty ʾšr ṣwty ʾwtm* "they have also transgressed my covenant which I commanded them"; Josh 23:16 *ʿbrkm ʾt bryt YHWH ʾlhykm ʾšr ṣwh ʾtkm* "if you transgress the covenant of YHWH your God, which he commanded you"; and Judg 2:20 *yʿn ʾšr ʿbrw hgwy hzh ʾt bryt ʾšr ṣwyty ʾt ʾbwtm* "since that nation has transgressed the covenant that I commanded their fathers" (SBH1).

[54] Compare: Num 19:2 *zʾt ḥqt htwrh ʾšr ṣwh YHWH* "rule of the law that YHWH has commanded" (SBH1); and 1 Chr 16:40 *wlkl hktwb btwrt YHWH ʾšr ṣwh* "according to all that is written in the law of YHWH which he commanded Israel"; Neh 8:1 *ʾt spr twrt mšh ʾšr ṣwh YHWH ʾt yśrʾl* "the book of the law of Moses, which YHWH had commanded Israel"; Neh 8:14 *wymṣʾw ktwb btwrh ʾšr ṣwh YHWH byd mšh* "they found it written in the law that YHWH had commanded by Moses" (LBH1).

[55] Compare Num 30:17 *ʾlh hḥqym ʾšr ṣwh YHWH ʾt mšh* "these are the statutes which YHWH commanded Moses" (SBH1).

(0/2, qiṭṭel). As Levine rightly pointed out, none of the terms mentioned above express inherently the idea of authority,[56] whereas *miṣwọt* can be accounted for as a full nominal lexicalization of it.

Among the heads governing *miṣwọt* within verbal phrases, the verbs for hearing, such as *šmʿ ʾel/ʾet* and *ʾzn* (H/1 hiqṭil) occupy a prominent position, alongside of the obvious *šmr* 'to keep,' *ʿśh* 'to do,' 'to put into practice,' and *ʿzb* 'to abandon,' pointing to the idea of compliance or non-compliance with the commandments.

One frequent construction is *šmʿ ʾl* 'to consent,' 'to listen to.'[57] Scholars have devoted special attention to the construction *šmʿ b*, especially to the text type *šamaʿ bəqọl* 'to obey.'[58]

Regarding the construction *šmʿ ʾl*, Arambarri has observed that it expresses 'approval, consent, acceptance, receipt, consent,' or, in the negated form, 'refusal.'[59] He has pointed out, moreover, that the meaning 'obey' turns out to be largely context-dependent, since it arises only under specific circumstances, namely when the approval is made binding on the basis of social or religious relations. The systematic analysis of the distribution of *šmʿ ʾl* in SBH1 and LBH1 reveals that the action described by the construction applies in particular to a kind of consent carried out freely, by people whose obedience does not derive from a bond of subordination to a person in control but rather from a personal conviction or resolution. In this pattern of usage, the indirect complement governed by the verb normally points to a person who has previously made a request or a demand. In the majority of cases, the persons to whom the subject of the verb consents are not in a position of control with respect to his or her will. This type of obedience appears to be based on the persuasion that the requested action is convenient. Many examples can be found in the historical-narrative language: Abraham accepts the terms of Ephron (*wyśmʿ ʾbrhm ʾl ʿprwn*) in the negotiations for the purchase of land (Gen 23:16); the sons of Jacob try to convince Shechem and his father Ḥămor (*wʾm lʾ tśmʿw ʾlynw*) to be circumcised (Gen 34:17); Pharaoh repeatedly refuses to consent to Moses' requests (Exod 6:30; 7:4.13.22; 8:11.15; 9:12; 11:9); Ben-Hadad, king of Aram is persuaded by Asa king of Judah (*wyśmʿ bn hdd ʾl hmlk ʾs*) to enter into alliance with him (1 Kgs 15:20); king Ahasuerus's attendants fail to convince Mordecai (*wlʾ šmʿ ʾlyhm*) to pay tribute to Aman (Esth 3:4).[60] When the indirect complement refers to a person with authority over the subject, it is normally a family relationship between parents (both mother and father) and children (Gen 28:7; 49:2; Deut 21:18). The role of king Solomon towards the people can be included in this framework (1 Chr 29:23). God is convinced by those who invoke him (Gen 30:17.22; Exod 22:23; Deut 3:26; 9:19; 17:12; 1 Kgs 8:52; 2 Kgs 13:4) especially though prayers and petitions.[61]

[56] According to Levine, the idea of authority is somehow superimposed on the core meaning of these terms, and often justified by other co-occurrent elements: "The *mišpāṭ* should be followed because it represents the accepted standard of justice … the *ḥōq* should be followed because someone with authority has written or promulgated it … the *torâ* should be followed because it has been presented or shown to someone … the word *miṣwâ* is authoritative in and of itself"; see Levine, *TDOT* 8: 506.

[57] See Deut 11:13.27.28; 28:13; for the text type *šamaʿ ʾel*, see *DCH* 8: 461.

[58] For the text type *šamaʿ bəqọl*, see Udo Rüterswörden, "עמש," *TDOT* 15: 265-266.

[59] See also Arambarri 1990: 154.

[60] Possibly the servants were genuinely concerned for Mordechai's safety in chiding him in a friendly way; see Moore 1971: 37.

[61] See *lšmʿ ʾl hrnh wʾl htplh* (1 Kgs 8:28, 29); *wšmʿt ʾl tḥnt ʿbdk* (1 Kgs 8:30); *lšmʿ ʾl tplt ʿbdk* (Neh 1:6).

In some of the passages quoted above[62] *miṣwọt* occurs as an indirect complement of *šmʿ ʾl*.[63] This construction is attested fifteen times in Deuteronomy,[64] with all the range of uses described.[65] Its distribution suggests that the reading that fits better corresponds to 'being persuaded to do something' rather than 'obeying as a subordinate.' Moreover, in hortatory discourse arguments in favor of obedience are mentioned on regular basis, that is, the reasons why it is convenient that the commandments are kept. This fact suggests that the kind of obedience expressed by the construction *šmʿ ʾl* needs to rely on adhesion resulting from a conscious conviction.

1.3 Divine standing order

One example in my database attests the usage of *miṣwà* for a standing order originating with God and imparted to a prophet acting as his attendant:

(19) 1 Kgs 13: 21-22
wyqrʾ ʾl ʾyš hʾlhym ʾšr bʾ myhwdh lʾmr kh ʾmr YHWH yʿn ky mryt py YHWH wlʾ šmrt ʾt hmṣwh ʾšr ṣwk YHWH ʾlhyk (22) *wtšb wtʾkl lḥm wtšt mym bmqwm ʾšr dbr ʾlyk ʾl tʾkl lḥm wʾl tšt mym lʾ tbwʾ nbltk ʾl qbr ʾbtyk*
'He (the old prophet living in Bethel) cried to the man of God who came from Judah, 'Thus says YHWH, because you have disobeyed the word of YHWH, and have not kept *the command*[66] which YHWH your God commanded you, (22) but have come back, and have eaten bread and drunk water in the place of which he said to you, Eat no bread, and drink no water; your body shall not come to the tomb of your fathers." (RSV)

The divine standing order to which this text refers is formulated for the first time in v. 9:

(20) 1 Kgs 13:9
ky kn ṣwh ʾty bdbr YHWH lʾmr lʾ tʾkl lḥm wlʾ tšth mym wlʾ tšwb bdrk ʾšr hlkt
'For so was it commanded me by the word of YHWH, saying, 'You shall neither eat bread, nor drink water, nor return by the way that you came." (RSV)

[62] See Deut 11:13; 11:27; 28:13.

[63] The same holds true for the combination *ḥqym wmšptym*; see, for exmple Deut 4:1 *wʿth yśrʾl šmʿ ʾl hḥqym wʾl mšptym ʾšr ʾnky mlmd ʾtkm lʿśwt lmʿn thyw wbʾtm wyrštm ʾt hʾrṣ ʾšr YHWH ʾlhy ʾbtykm ntn lkm* "and now, O Israel, listen to the statutes and the rules that I am teaching you, and do them, that you may live, and go in and take possession of the land that YHWH, the God of your fathers, is giving you."

[64] The textual type *šmʿ ʾt* is also admitted, it combines with *ḥqym* (Deut 4:6), *dbry* (Deut 4:10), *ḥqym wmšptym* (Deut 5:1), *kl ʾšr yʾmr YHWH ʾlhynw* (Deut 5:27), *mšptym* (Deut 7:12), *dbrym* (Deut 12:28, and 29:18), and particularly *qwl* (Deut 1:34; 4:36; 5:23.24.25; 5:28; 18:16; 26:7). The construction *šmʿ b* occurs only with *qwl* (Deut 1:45, with God as subject), and mostly with *qwl YHWH* (Deut 4:30; 8:20; 9:23; 13:5.19; 15:5; 21:18.20; 26:14.17; 27:10; 28:1.2.15.45.62; 30:2.8.10.20).

[65] God can consent to the someone's requests or not (Deut 3:26; 9:19; 10:10; 23:6); one shall not be persuaded to idolatry by the enticing speeches of a prophet or a seer, or a brother, a son, a daughter, a beloved wife, or friend (*dbry hnbyʾ hhwʾ ʾw ʾl ḥwlm ḥḥlwm hhwʾ*, 13:4.9); one must obey the priest and the judge (Deut 17:12); the nations listen to fortune-tellers and to diviners (18:14); a son must listen to the voice of the father and the mother (*šmʿ bqwl ʾbyw wbqwl ʾmw*), and obey them (*yšmʿ ʾlyhm*, 21:18); Moses' endorsement of Joshua places him in the position of being obeyed by the people (34:9); and finally, divine *mṣwt* must be obeyed (4:1; 11:13; 11:27.28; 28:1).

[66] RSV translates "the commandment.

This is not an absolute prohibition but a contingent command, valid in the situation represented by the narrative.

2. Expression of human authority

I have shown above that the reference to the divine origin of the command is not fully lexicalized in the substantive *miṣwà*. In fact, in Biblical narrative *miṣwà* applies also to binding instructions given by authorities to people in a subordinate position; the noun occurs particularly in the framework of royal and military commands.

Obedience in this case does not imply an act of a free decision but it appears as a duty, an obligation, or a responsability. With reference to its effect, such a command may be valid under given circumstances or retained irrespective of changing conditions. In the latter case, the order is a directive made known publicly by kings (David, Solomon, Hezekiah, Joash, Josiah, and Ahasuerus) [67] or officers in charge (*śarim*), which is binding on all people under their command, and intended to enforce a policy or a procedure. Such commands may be issued orally or may imply a written form. The typical structuring of this reading is *miṣwat-* (singular construct plus governed noun pointing to a human authority).

2.1 Standing order

Many examples can be found in SBH1of specific commands issued orally under certain circumstances:

(21) 2 Kgs 18:36
whḥryšw hʿm wlʾ ʿnw ʾtw dbr ky mṣwt hmlk hyʾ lʾmr lʾ tʿnhw
'But the people were silent and answered him not a word, for *the king's command* was, 'Do not answer him."[68]

The reading of *miṣwà* in 1 Kings 2:43 must be included in this group. In the narrative, one of the first acts of Solomon as a king is to enjoin Shimei to reside in Jerusalem, depriving him on pain of death of the freedom to move.[69] This action is expressed by the verb *ʿawad* (H/1 hiqṭil) 'to admonish,' 'to warn'[70] (1 Kgs 2:42). It is remarkable that Solomon had Shimei swear by YHWH, suggesting either that the royal order (*hammiṣwà ʾăšer ṣiwwiti ʿaleḵa*, v. 2:43) was not such a peremptory constraint *per se* or that Solomon felt himself not sufficiently established in his authority.[71]

In LBH1 a sentence of death by stoning issued by king Joash against the prophet Zechariah constitutes *miṣwat hammeleḵ*, which is immediately and publicly carried out:

[67] See Isa 36:21; 2 Kgs 18:36 (SBH1); and Esth 3:3; 2 Chr 8:14.15; 24:21; 29:15.25; 30:6.12; 35:10.15.16; Neh 11:23; 12:24.45 (LBH1).

[68] Parallel to Isa 36:21 *whḥryšw wlʾ ʿnw ʾtw dbr ky mṣwt hmlk hyʾ lʾmr lʾ tʿnhw.*

[69] The Solomon's command is expressed through a series of directive verbal forms: *bnh … wyšbt … wlʾ tṣʾ* (see 1 Kgs 2:36).

[70] For the meaning of the denominative verb *ʿwd* H/1 hiqṭil, see *HALOT*, 6843.

[71] See 1 Kgs 2:42 *hlwʾ hšbʿtyk bYHWH wʾʿd bk lʾmr* "did I not make you swear by YHWH and solemnly warn you."

(22) 2 Chr 24:21
wyqšrw ʿlyw wyrgmhw ʾbn bmṣwt hmlk bḥṣr byt YHWH
'They conspired against him, and *by command of the king* they stoned him with stones in the court of the house of YHWH.' (RSV)

In Esther 3:3, the royal order consists of bowing before Haman the Agagite and paying homage to him (cf. v. 2). In Qoheleth 8:5, the obedience to the king's command is encouraged as it provides prosperity and success; in this passage the term occurs in absolute case (*hammiṣwà*), being coreferential to the previous expressions *pi meleḵ* 'king's command' (v. 2), and *dəḇar meleḵ* 'king's word' (v. 4).

2.2 Royal regulation

Frequently, and increasingly in later layers of language, the term refers to more complex regulations, typically issued by kings and intended to enforce a policy, with special reference to the religious domain and cultic matters. In 2 Chronicles, David is depicted as the prime example of the reformer who organizes the clergy.[72] Moreover, the usage of regulation formulas punctuates the description of the celebrations of Passover at Jerusalem during the kingdoms of Hezekiah (2 Chr 30) and Josiah (2 Chr 35: 1-18):

- *bəmiṣwat Dawid wəgad ḥozẹr hammeleḵ wənatan hannaḇiʾ* 'according to the command of David and of Gad the king's seer and of Nathan the prophet' (2 Chr 29:25)
- *kəmiṣwat Dawid* 'according to David's command' (2 Chr 35:15)
- *kətọrat Mọšè ʾiš kətọrat Mọšè ʾiš hå'Ělọhim* 'according to the teaching of Moses, the man of God' (2 Chr 30:16)
- *biḵtaḇ Dawid meleḵ Yiśraʾel uḇəmiḵtaḇ Šəlọmọ̀ bənọ* 'as prescribed in the writing of David king of Israel and the document of Solomon his son' (2 Chr 35:4)[73]
- *kəmiṣwat hammeleḵ* 'according the king's command' (2 Chr 29:15; 35:10)
- *kakkatuḇ bəseṗer Mọšè* 'as it is written in the book of Moses' (2 Chr 35:12)
- *kəmiṣwat hammeleḵ Yọšiyyahu* 'according to the command of king Josiah' (2 Chr 35:16).

According to de Vries, the Chronicler does not dispute the prime authority of Moses as cult founder,[74] but he is concerned about establishing David's authority, especially "because there was no clear consensus in postexilic Israel about a continuing role for David's successors."[75] The Chronicler intends to express the idea that the Davidic line's duty was simply to carry out the regulations that David laid down.

[72] See 2 Chr 8:14.15. Interestingly enough, there is an overlap between the expression *miṣwat Dawid* and *miṣwat Mọšè* in this particular usage (compare 2 Chr 8:13.14). It is important to highlight the summarizing effect of the authorization formula *kĕmiṣwat Mọšè*, that turns out to be put in operation when some specific mode or repository of revelation needs to be mentioned (2 Chr 8:13). The reference is equal to the entire corpus of laws regulating the sacred festivals in this case; see de Vries 1988: 621.

[73] The Chronicler appeals to a document concerning the Levitical preparation of *Pesaḥ* written by David and then actualized by Solomon.

[74] According to Williamson, "we may confidently assert that the Chronicler had the Pentateuch before him in its final and completed form"; see Williamson 1976: 361.

[75] See de Vries 1988: 631-632; according to the Chronicler's understanding, moreover, inspiration was not limited to figures that were commonly identified as "prophets"; direct communication with God is ascribed also to the founding kings of the Davidic dynasty; see Japhet 1993: 46; Petersen 1977.

Such regulation formulas occur also in Nehemiah with a comparable function; they are put into operation to highlight that the legitimate performance of liturgical duties, established by David, was faithfully implemented by Solomon and his descendants:

- *bəmiṣwat Dawid ʾiš haʾĔlọhim* 'according to the command of David the man of God' (Neh 12:24).
- *kəmiṣwat Dawid Šəlọmọ̀ bənọ* 'according to the command of David, and of Solomon his son' (Neh 12:45).[76]

The Chronicler depicts king Hezekiah with special emphasis as the champion of the restoration of the cult at the Jerusalem Temple.[77] A royal decree issued by him and dispatched by his messengers is called *miṣwat hammeleḵ* (vv. 6-9):

> (23) 2 Chr 30:6.8
> *wylkw hrṣym bʾgrwt myd hmlk wśryw bkl yśrʾl wyhwdh wkmṣwt hmlk lʾmr bny yśrʾl šwbw ʾl YHWH ʾlhy ʾbrhm yṣḥq wyśrʾl wyšb ʾl hplyth hnšʾrt lkm mkp mlky ʾšwr …* (8) *ʿth ʾl tqšw ʿrpkm kʾbwtykm tnw yd lYHWH wbʾw lmqdšw ʾšr hqdyš lʿwlm wʿbdw ʾt YHWH ʾlhykm wyšb mkm ḥrwn ʿpw*
> 'So couriers went throughout all Israel and Judah with letters from the king and his princes, *according to the command of the king* which was: 'O people of Israel, return to YHWH, the God of Abraham, Isaac, and Israel, that he may turn again to the remnant of you who have escaped from the hand of the kings of Assyria … (8) Do not now be stiff-necked as your fathers were but yield yourselves to YHWH and come to his sanctuary, which he has consecrated forever, and serve YHWH your God, that his fierce anger may turn away from you.'

Hezekiah's *miṣwà* reported in this text turns out to be tantamount to a strong appeal to the people of the old Northern Kingdom, which by that time had become a province of Assyria, to join their brothers at Jerusalem for celebrating *Pesaḥ* together. The text of the decree must be regarded as a heartfelt exhortation rather than an order, owing to the fact that the people of the North were not actually his real subjects. Verse 6 contains a kind of conflation: the letters (*ʾiggərọt*) are said to be "from the king and his princes" (*miyyad hammeleḵ wəśarayw*); the message dispatched, nevertheless, is spoken according "the king's command" (*uḵəmiṣwat hammeleḵ*). Japhet thinks that this fact reflects the Chronicler's attitude towards the kingdom, and the reported decree itself must be regarded as "an outstanding example of the Chronicler's literary methods and theological positions."[78] Namely, on the one hand, the narrative highlights the collegial nature of the kingdom's administration by mentioning the princes; on the other hand, it is meant to restate Hezekiah's function and authority.

[76] See Blenkinsopp 1988: 350.

[77] Hezekiah's restoration of the temple in Chronicles (missing in the books of Kings) makes him another temple builder, long with David and Solomon, and his celebration of Passover (also missing in Kings) is treated at length (2 Chr 30); see French 2017: 148-154.

[78] See Japhet 1993: 941.

2.3 Will

Besides kings, fathers as well can transmit instructions to sons as an expression of their will,[79] especially before death. Such instructions are regarded by the posterity as a *miṣwà*:

> (24) Jer 35:16
> *ky hqymw bny yhwndb bn rkb ʾt mṣwt ʾbyhm ʾšr ṣwm whʿm hzh lʾ šmʿw ʾly*
> 'Indeed, the sons of Jonadab the son of Rechab have carried out *the command of their father* which he commanded them, but this people has not listened to me.'

Jonadab's command to his sons consists actually in not drinking wine (*lblty štwt yyn*, v. 14), and it can be conceptualized either as uniplex entity,[80] or as a multiplex entity.[81] Remarkably, the verb *qwm* (H/1 hiqṭil) is used for the action of complying with the father's instructions in this passage[82] instead of the more obvious *šāmaʿ ʾel* or *ʿāśà*.

3. Peripheral sense-nodules: polysemy as a window on diachronic change

A group of later attestations witnesses a remarkable development in the pattern of usage of *miṣwà*, which have a strong impact on its reading and may be regarded as signals of an ongoing change in its semantics.

As I have observed before, a typical feature of deverbal nouns is to inherit the valency frame slots of their source verbs.[83] In BH *ṣwh* (0/2 qiṭṭel) exhibits a very large range of different syntactic constructions,[84] which can be traced back to two main valency frames: on the one hand, it is used as a three-argument verb, requiring an *actor* (or agent), viz. the human or divine person in control who performs the action of commanding; an *addressee*, viz. the person to whom the order is directed; and a *patient*, vz. the action to be performed that is the object of the command. On the other hand, *ṣwh* (0/2 qiṭṭel) is attested as a two-argument verb, with the surface deletion of the object, meaning 'to give orders'.[85]

Concerning nouns derived by syntactic derivation, participants are normally expressed in BH by governed nouns or pronominal suffixes. The nominal complements of *miṣwà* point exclusively to the agent in SBH1, namely to the subject provided of animacy who issues the command. Expressions such *miṣwat YHWH, miṣwat hammelek* or *miṣwat Yĕhoṇadaḇ* must be thus read as "the command which YHWH/the king/Jonadab had issued." This compact trend will undergo some variation in LBH1. A few examples of the usage of the term in the book of Nehemiah are particularly telling. They attest a remarkable shift in the arguments expressed on the surface as governed nouns, affecting the reading of *miṣwà* in the direction described below.

[79] The verb *ṣwh* as well can take on a similar reading, compare Gen 49:29.33; 50:16; Isa 38:1.

[80] See MT *miṣwat ʾăḇihem* (Jer 35:14, 16) and *miṣwat Yĕhoṇadaḇ* (35:18).

[81] See MT *ʾet kol miṣwôtāyw* (Jer 35:18).

[82] Compare Jer 35:14 *hwqm ʾt dbry yhwndb bn rkb* "the commands of Jonadab son of Rechab have been fulfilled"; for the meaning of *qwm* H/1 hiqṭil, see *HALOT*, 8302 "to take out," "to keep," when the verb selects as objects *daḇar*, *neḏer, bərit, šəbuʿà*.

[83] See Panevová 2014: 7-11.

[84] Clines lists 23 of them, see *DCH* 7: 94-102.

[85] See, for example, Gen 49:33 *wykl yʿqb lṣwt ʾt bnyw wyʾsp rglyw ʾl hmṭh* "when Jacob finished commanding (giving instructions to) his sons, he drew up his feet into the bed."

3.1 Obligation, duty

Semantically speaking, the shift from 'command' to 'obligation' or 'duty,' verifiable in some late uses of the noun *miṣwà*, is quite understandable and can be accounted for in terms of *converseness*. The category of converseness, borrowed from the science of symbolic logic, is used by semanticians to name a subclass of oppositeness implying a mirror-image relation between a pair of lexical items, called thus *converses*. Cruse describes converses as *relational opposites*,[86] which refer to the same relationship from reversed points of view. Converses may imply reciprocity (as *friend* or *mate*)[87] or asymmetry (as *doctor* vs. *patient* or *teach* vs. *learn*).

Operations of permuting the arguments of a pair of converses can help appreciate the sense-relation at stake; the sentence "Tom *sells* his car to Sam," for example, entails logically the sentence "Sam *buys* Tom's car"; that being the case, we can safely consider *sell* and *buy* converses. If we apply such a test to the sense-nodules ascribable to *miṣwà*, it is clear that an expression like "the king's command to the people" logically entails its reversed counterpart "the duty of the people towards the king"; in the first wording the action is regarded from the point of view of its actor (the king), in the latter case from the point of view of its recipient (the people). I can affirm, therefore, that the meaning "command" underwent a conceptual re-analysis developing the converse sense, of 'duty.'[88] Concerning *miṣwà*, such a shift occurs on the level of the semantic micro-structure of the noun, yielding the phenomenon of auto-converseness. LBH1 mirrors the beginnings of this semantic development. A clear example of this converse sense-nodule is attested in the following passage:

> (25) Neh 10:33
> *wh'mdnw 'lynw mṣwt ltt 'lynw šlšyt hšql bšnh l'bdt byt 'lhynw*
> 'We also lay upon ourselves *the obligation* to charge ourselves yearly with the third part of a shekel for the service of the house of our God.'[89]

The shift in perspective is further emphasized by the verbal selector *'md* (H/1 hiqṭil) *'l* 'to lay upon.'[90] It is important to point out that the sense-nodule 'obligation' would play a central role for the further semantic development of the term in post-biblical layers of Hebrew language, mostly within the halakhic-rabbinic discourse tradition.

4. Remarks on the semantic range of miṣwà *in the Mishnà*

The increase in the frequency of the lexeme *miṣwà* already observed within LBH1 is further witnessed by the Mishnà, as the following table shows:

[86] See Cruse 1986: 231; see also König 2011: 331.

[87] In these cases, we can speak of auto-converseness.

[88] See Rainer 2005.

[89] Among modern translation, some mirror the shift here highlighted, compare: "we have laid upon ourselves obligations" (RSV; NJPS); "we also placed ourselves under obligation" (NASB); and "we recognize the following obligations" (NJB); others opt for circumlocutory phrases as: "we hereby undertake the duty" (NEB); and "we assume the responsibility for carrying out the commands" (NIV).

[90] See *DCH* 6: 474, 8b.

	Singular	*Plural*	*TOT*
SBH1	1.72	3.54	5.27
LBH1	5.16	3.75	8.91
MH	7.36	1.45	8.81

Table 1 – Frequency ratio (normalized per 10,000 words)[91]

The readings 'religious duty,' and 'religious requirement' ends up prevailing in Mishnaic Hebrew.[92] This phenomenon is evident above all in the productive discourse, while in quotations from the Scripture, the SBH main meaning 'commandment' (in the plural) still stands, albeit considerably marginal.[93] For the present survey, I have collected the following examples, emerging from the syntagmatic analysis,[94] which clearly demonstrate how the auto-converse reading already found in LBH1 texts has become fully lexicalized in the Mishnà as the most salient meaning of the *miṣwà*.

(26) *m.Yebam*. 4.5
Mṣwh bgdwl lybm lʾ rṣh mhlkyn ʿl kl hʾḥyn lʾ rṣw ḥwzrym ʾṣl hgdwl wʾwmʾ lw ʿlyk mṣwh ʾw ḥlwṣ ʾw yybm
Tlh bqṭwn ʿd šygdyl wbgdwl ʿd šybwʾ mmdynt hym wbḥrš wbšwṭh ʾyn šwmʿym lw ʾlʾ ʾwmryn lw ʿlyk mṣwh
'A. It is *the duty of the oldest (surviving) brother* to enter into levirate marriage. B. If he did not want to do so, they pass in turn to all the other brothers. C. If they all did not want to do so, they go back to the oldest and say to him, 'Yours is the duty!' (Neusner)

This passage is quite remarkable. The precepts concerning the levirate marriage, formulated in Deut 25: 5-10, do not include any special mention of the oldest brother as the specific recipient of the commandment of marrying the brother's childless widow. This context seems to trigger the reading 'the proper/perfect way of fulfilling a religious duty'[95] which could be related to the further development of the term in rabbinic Hebrew pointing to the sense-nodule 'meritorious deed'.[96]

(27) *m.Yebam*. 12.6
Wnqrʾ šmw byśʾ byt ḥlwṣ hnʿl mṣwh bdyynyn wlʾ mṣwh btlmydym rʾ yhwdh ʾwʾ mṣwh ʿl kl hʿwmdym šm lwmr ḥlwṣ hnʿl ḥlwṣ hnʿl ḥlwṣ hnʿl
L. And his name shall be called in Israel: 'The house of him who has had his shoe removed' (Deut 25:9)—it is *the duty of the judges*, and not *the duty of the disciples* [so to name him]. M. R. Judah says, 'It is *the duty of all bystanders* to

[91] The noun occurs 158 times in MH, 132 in the singular and only 26 time in the plural, see appendix.
[92] See Jastrow, *s.v.*; the noun also became an antonym of *rəšut* in its reading "optional deed", "act spread from self-determination", cf. *m.Pes*. 6.2; *m.Makk*. 2.7.12; *m.Betz*. 5.2.
[93] See, for example, *m.Qid*. 4.14; *m.Hor*. 3.3; *m.Men*. 9.7.
[94] See Appendix.
[95] Cf. also *m.Yebam*. 12.6: *mṣwt ḥlyṣh* "the proper way to carry out the rite of *ḥaliṣà* (is as follows)".
[96] The noun *miṣwà* will pass to designate any act of human kindness, such as the burial of the body of an unknown person (*męt miṣwà*, see m.Naz. 6.5; 7.1).

say, 'The man whose shoe has been removed! The man whose shoe has been removed! The man whose shoe has been removed!". (Neusner)

This passage refers to a particular aspect of the rite of *ḥaliṣà*, by which the surviving brother could evade the obligation of the levirate marriage. The ceremony involves the widow making a declaration, taking off a shoe of her brother-in-law, and spitting on the floor. This rite also implies the stigmatization of the refusal to fulfil this religious duty; the person in question will be given a title with which it will be recognized by the whole community. A quite relevant aspect of this passage is that the noun *miṣwà* governs a Pph with the preposition *ʿl* indicating the subject on whom the duty falls. This is the same construal that is found in Neh 10:33 and 11:23, which has no other parallel in BH.

(28) *m.Yoma* 8.4
htynwqwt ʾyn mʿnyn ʾwtm bywm hkypwrym ʾbl mḥnkyn ʾwtm qwdm (l)šnh wqwdm štym bšbyl šyhw rgylym lmṣwt
'A. As to children, they do not impose a fast on them on the Day of Kippur. B. But they educate them a year or two in advance, so that they will be used to doing *the religious duties*.' (Neusner)

(29) *m.ʾAḇot 2.1*, 2b
zhyr bmṣwh qlh kmṣwh ḥmwrh šʾyn ʾth ywdʿ mtn śkrn šlmṣwt
'Be meticulous *in a small religious duty* as in a large one, for you do not know what sort of reward is coming for any *of the various religious duties*.' (Neusner)

This semantic shift is also accompanied by a significant change in the syntagmatic pattern of usage of the word, namely in its valency frame. In the repeated discourse (viz. in quotations from the Scripture), the genitive or the pronominal suffix governed by *miṣwà* encodes exclusively the actor who issues the commandment (mostly YHWH); in the productive discourse, on the other hand, the genitive complement points normally either to the subject of the obligation or to its recipient. The following examples show both cases respectively:

(30) *m.Suk.* 2.7
my šhyh rʾšw wrwbw bswkh wšwlḥnw btwk hbyt, byt Šmy pwslyn wbyt Hll mkšyryn (…) *ʾmrw lhm byt Šmy mšm rʾyh ʾp hn ʾmrw lwʾm kk hyth (lw) nwhg lʾ qyymth mṣwt sukh mymyk*
'A. He whose head and the greater part of whose body are in the *sukkah*, but whose table is in the house. B. The House of Šammai declare invalid. C. And the House of Hillel declare valid. (…) E. Said the House of Šammai to them, 'Is there proof from that story? But in point of fact they did say to him, 'If this is how you act, you have never in your whole life fulfilled *the religious requirement of dwelling in a sukkà*!" (Neusner)

(31) *m. Pes.* 3.7, 1c
Hhwlk lšḥwṭ ʾt psḥw wlmwl ʾt bnw wlʾwkl sʿwdt ʾyrwsym bbyt ḥmyr wnzkr šyš lw ḥmṣ btwk hbyt ʾm ykwl lḥzwr wlbʿr wlḥzwr bmṣwtw yḥzwrn
'He who goes to slaughter his *Pesaḥ* lamb, to circumcise his son, or to eat the betrothal meal at his father-in-law's house, and remembers that he has

left some leaven in his house, if he can go back and remove it and go on to do his religious duty (*lmṣwtw*), let him go back and remove it.' (Neusner)

Concluding remarks

The data collected for this paper have shown that from Biblical to Mishnaic Hebrew the noun *miṣwà* underwent a semantic shifting which can be accounted for in terms of *auto-converseness* from the meaning of 'commandment', mostly applying to commandment originated in God and disclosed through the Mosaic teaching, to the meaning of 'religious duty'.

This semantic shift, however, must not be regarded as an innovation peculiar to the Rabbinic Hebrew; its arising can be traced back to the Late Biblical Historical-narrative language. The lexeme *miṣwà* is a good example of how a peripheral sense of a polysemic word in a given linguistic stratum can be the vector of an extensive diachronic change.

Abbreviations

BH	Biblical Hebrew
SBH	Standard Biblical Hebrew
LBH	Late Biblical Hebrew
ABH	Archaic Biblical Hebrew
SBH1	Standard Biblical Hebrew – Historical-narrative language
SBH2	Standard Biblical Hebrew – Poetical language
SBH3	Standard Biblical Hebrew – Language of Hosea
LBH1	Late Biblical Hebrew – Historical-narrative language
LBH2	Late Biblical Hebrew – Poetical language
MH	Mishnaic Hebrew
BDB	Brown, Francis, S. R. Driver, and Charles A. Briggs, *A Hebrew and English Lexicon of the Old Testament with an appendix containing the biblical Aramaic*. CD-ROM edition. Oxford: Clarendon, 1906.
DCH	Clines, David J. A., ed. *Dictionary of Classical Hebrew*. 9 vols. Sheffield: Sheffield Phoenix Press, 1993–2014.
HALOT	Koehler, Ludwig, and Walter Baumgartner. *The Hebrew and Aramaic Lexicon of the Old Testament: The New Koehler-Baumgartner in English*. Subsequently revised by Walter Baumgartner and Johann Jakob Stamm with Assistance from Benedikt Hartmann, Ze'ev Ben-Hayyim, Eduard Yechezkel Kutscher, and Philippe Reymond. Translated and edited under the supervision of M. E. J. Richardson. CD-ROM edition. Leiden: Brill, 1994-2000.
NASB	*New American Standard Bible*.
NEB	*The New English Bible*.
NET	*The NET Bible*. Version 1.0. Biblical Studies Foundation.
NETS	*A New English Translation of the Septuagint*. Edited by Albert Pietersma and Benjamin G. Wright. Oxford: Oxford UP, 2007.
NIV	*The Holy Bible: New International Version*.
NJB	*New Jerusalem Bible*.
NJPS	*Tanakh: The Holy Scriptures: The New JPS Translation according to the Traditional Hebrew Text*.
NKJV	*New King James Version*.
RSV	*Revised Standard Version*.

References

Arambarri, Jesús. 1990. *Der Worstamm „hören" im Alten Testament. Semantik und Syntax eines Hebräischen Verbs*. Stuttgart: Verlag Katholisches Bibelwerk.

Avishur, Yitzhak. 1972. "Pairs of Synonymous Words in the Construct State and in Appositional Hendiadys in Biblical Hebrew." *Semitics* 2: 7-81.

Baroni, Marco. 2008. "Distributions in text." In *Corpus Linguistics. An International Handbook*, ed. by Anke Lüdeling, and Merja Kytö, vol. 1, 803-821. Handbücher zur Sprach- und Kommunikationswissenschaft 29.1. Berlin: Mouton de Gruyter.

Blenkinsopp, Joseph. 1988. *Ezra-Nehemiah*. Old Testament Library. London: SCM Press.

Cohen, Shaye J.D. 2007. "The Judaean Legal Tradition and the Halakhah of the Mishnah". In *The Cambridge Companion to the Talmud and Rabbinic Literature*, ed. by Charlotte Elisheva Fonrobert, 121-143. Cambridge: Cambridge UP.

Cruse, D. Alan. 1986. *Lexical Semantics*, Cambridge Textbooks in Linguistics. Cambridge: Cambridge UP.

Cruse, D. Alan. 2000. "Aspects of the Micro-structure of Word Meaning." In *Polysemy. Theoretical and Computational Approaches*, ed. by Yael Ravin, and Claudia Leacock, 30-51. New York: Oxford UP.

De Vries, Simon J. 1988. "Moses and David as Cult Founders in Chronicles." *Journal of Biblical Literature* 107: 619-639.

Fox, Michael V. 2009. *Proverbs 10-31*. The Anchor Bible 18b. New Haven/London: Yale UP.

French, Blaire A. 2017. *Chronicles Through the Centuries*, Wiley Blackwell Bible Commentaries. Oxford: Wiley Blackwell.

Geeraerts, Dirk, and Gitte Kristiansen. 2019. "Variationist Linguistics." In *Cognitive Linguistics. A Survey of Linguistic Subfields*, ed. by Ewa Dąbrowska, and Dagmar Divjak, 133-158. Berlin-Boston, MA: De Gruyter.

Grondelaers, Stefan, Dirk Speelman, and Dirk Geeraerts. 2007. "Lexical Variation and Change." In *The Oxford Handbook of Cognitive Linguistics*, ed. by Dirk Geeraerts, and Hubert Cuyckens, 988-1011. Oxford: Oxford UP.

Houtman, Cornelis. 2000. *Exodus*, vol. 3. Leuven: Peeters.

Hurvitz, Avi. 1972. *The Transition Period in Biblical Hebrew, A Study in Post-Exilic Hebrew and Its Implications for the Dating of Psalms*. Jerusalem: Bialik Institute [Hebrew].

Hurvitz, Avi. 1995. "Continuity and Innovation in Biblical-Hebrew. The Case of Semantic Change in Post-exilic Writings." In *Studies in Ancient Hebrew Semantics*, ed. by Takamitsu Muraoka, 1-11. Abr-Nahrain: Supplement Series 4. Leuven: Peeters.

Japhet, Sara. 1993. *I & II Chronicles*. Old Testament Library. London: SCM Press.

Jastrow, Marcus. 1903. *A Dictionary of the Targumim, the Talmud Babli and Yerushalmi, and the Midrashic Literature*. 2 vols. New York, NY: G. P. Putnam's Sons.

Jenni, Ernst. 1968. *Das hebräische Pi'el. Syntaktisch-semasiologische Untersuchung einer Verbalform im Alten Testament*. Zürich: Evz Verlag.

Joüon, Paul, and Takamitsu Muraoka. 2006. *A Grammar of Biblical Hebrew*. Subsidia Biblica 27. Rome: Biblical Institute Press.

König, Ekkehard. 2011. "Reciprocals and Semantic Typology: Some Concluding Remarks." In *Reciprocals and Semantic Typology*, ed. by Nicholas Evans, Alice Gaby, Stephen C. Levinson, and Asifa Majid, 329-340. Typological Studies in Language 98. Amsterdam-Philadelphia, PA: John Benjamins.

Kuryłowicz, Jerzy. 1936. "Dérivation lexicale et dérivation syntactique." *Bulletin de la Société linguistique de Paris* 37: 79-92.

Kuntz, J. Kenneth. 2004. "Hendiadys as an Agent of Rhetorical Enrichment in Biblical Poetry, with Special Reference to Prophetic Discourse." In *God's Word for Our World*, ed. by Simon J. De Vries, Harold J. Ellens, Deborah L. Ellens, Rolf P. Knierim, and Isaac Kalimi, vol. 1, 114-135. New York: T and T Clark.

Levine, Baruch A. 2000. "הוצמ." In *Theological Dictionary of the Old Testament*, ed. by Gerhard Johannes Botterweck, Helmer Ringgren, Heinz-Josef Fabry, trans. by David E. Green, vol. 8, 505-514. Grand Rapids, MI: Eerdmans.

Liedke, Gerhard. 1997. "הוצ." In *Theological Dictionary of the Old Testament*, ed. by Ernst Jenni, and Claus Westermann, trans. by Mark E. Biddle, vol. 2, 1062-1066. Peabody, MA: Hendrickson Publishers.
Lillas-Schuil, Rosmari. 2006. "A Survey of Syntagms in the Hebrew Bible Classified as Hendiadys." In *Current Issues in the Analysis of Semitic Grammar and Lexicon*, ed. Lutz Edzard, and Jan Retsö, 79-99. Wiesbaden: Harrassowitz Verlag.
Moore, Carey A. 1971. *Esther*. The Anchor Bible 7b. Garden City, NY: Doubleday.
Netzer, Yael. 2013. "Quantifier." *Encyclopedia of Hebrew Language and Linguistics*, ed. by Geoffrey Khan, Shmuel Bolozky, Steven Fassberg, Gary A. Rendsburg, Aaron D. Rubin, Ora R. Schwarzwald, and Tamar Zewi, vol. 3, 311-315. Leiden-Boston, MA: Brill.
Neusner, Jacob. 1991. *The Mishnah: A New Translation*. London-New Haven, CT: Yale UP.
Panevová, Jarmila. 2014. "Contribution of Valency to the Analysis of Language." In *Noun Valency*, ed. by Olga Spevak, 1-17. Studies in Language Companion Series 158. Amsterdam-Philadelphia, PA: John Benjamins.
Petersen, David L. 1977. *Late Israelite Prophecy: Studies in Deutero-Prophetic Literature and in Chronicles*. Society of Biblical Literature Monograph Series. Missoula, MT: Scholars Press.
Propp, William H.C. 2006. *Exodus 19-40, The Anchor Bible 2a*. New Haven-London: Yale UP.
Rainer, Franz. 2005. "Semantic Change in Word Formation." *Linguistics* 43 (2): 415-441.
Rofé, Alexander. 2002. "The Book of Deuteronomy: A Summary." In *Deuteronomy*, *Issues and Interpretation*, 1-13. London-New York, NY: T and T Clark.
Sáenz-Badillos, Angel. 1993. *A History of the Hebrew Language*, English trans. by John Elwolde. Cambridge: Cambridge UP.
Strack, Hermann L., and Günter Stemberger. 1992. *Introduction to the Talmud and Midrash*, English trans. by Markus Bockmuehl. Philadelphia, PA: Fortress Press.
Vergari, Romina. 2021. *Toward a Contrastive Semantics of the Biblical Lexicon. The Nouns of "Rule and Regulation" in Hebrew Historical-narrative Language and their Greek Equivalents*. Topografie immateriali 2. Firenze: SEF.
Vivian, Angelo. 1978. *I campi lessicali della separazione nell'ebraico biblico, di Qumran e della Mishna: ovvero, applicabilita della teoria dei campi lessicali all'ebraico*. Quaderni di Semitistica 4. Florence: Istituto di linguistica e di lingue orientali.
Weinfeld, Moshe. 1972. *Deuteronomy and the Deuteronomistic School*. Oxford: Clarendon Press.
Weinfeld, Moshe. 1991. *Deuteronomy 1-11*. The Anchor Bible 5. Garden City, NY: Doubleday.
Weiser, Artur. 1962. *The Psalms*. Old Testament Library. London: SCM Press.
Williamson, Hugh G.M. 1976. "The Accession of Solomon in the Books of Chronicles." *Vetus Testamentum* 26: 351-361.
Zatelli, Ida. 1978. *Il campo lessicale degli aggettivi di purità in ebraico biblico*, Quaderni di Semitistica 7. Florence: Istituto di linguistica e di lingue orientali.
Zatelli, Ida. 1995. "Functional Languages and Their Importance to the Semantics of Ancient Hebrew." In *Studies in Ancient Hebrew Semantics*, ed. by Takamitsu Muraoka, 55-63. Abr-Nahrain: Supplement Series 4. Louvain: Peeters.
Zatelli, Ida. 2004. "The Study of the Ancient Hebrew Lexicon. Application of the Concepts of Lexical Field and Functional Language." *KUSATU* 5: 129-159.

Appendix: Distribution of miṣwà in MH

The noun *mṣwh* occurs 158 times, according to the following distribution:

TOT	**MH**
bmṣwh	1
bmṣwtw	1
hmṣwh	3
hmṣwt	6
hmṣwwt	3
kmṣwh	1
kmṣwth	1
kmṣwtw	5
kmṣwwtn	2
kmṣwwtw	10
lmṣwh	2
lmṣwt	3
lmṣwwt	2
mṣwh	54
mṣwht	1
mṣwt	32
mṣwtn	5
mṣwtw	1
mṣwwh	1
mṣwwt	10
mṣwwtk	1
mṣwwtn	3
mṣwwtw	1
šlmṣwt	1
šmṣwh	1
šmṣwwt	1
šmṣwwtw	3
wmṣwh	1
wmṣwt	1
wmṣwwt	1
TOT	**158**

Singular forms (132)
m.Ab. 2.1(x3); 4.2(x5); 4.11
m.Bek. 1.7(x11); 8.6(x2)
m.Ber. 1.1(x3)
m.Betz. 5.2(x2)
m.BQ. 2.10(x2)
m.Ed. 5.3
m.Hor. 2.4(x2); 2.6
m.Hul. 7.2; 12.4(x2)
m.Ket. 4.3
m.Makk. 1.7(x2); 2.7(x2); 3.4; 3.15(x2)
m.Meg. 2.6(x3); 4.8
m.Meg. 1.3(x5); 3.7; 9.7; 10.2; 10.4; 10.9(x3)
m.Naz. 6.5(x2); 7.1
m.Neg. 14.4(x4); 14.10
m.Nid. 10.1
m.Par. 11.9(x2); 12.4
m.Pes. 3.7; 6.2(x3); 10.3(x2)
m.Qid. 1.7(x3); 1.9(x3)
m.RH. 3.3(x2)
m.San. 7.1(x5); 8.1
m.Shebu. 2.3(x2); 3.6(x4); 3.8
m.Sot. 8.7(x2)
m.Suk. 2.4; 2.7; 4.4(x2)
m.Yeb. 2.3(x2); 2.8; 3.2; 3.4; 4.5(x3); 5.2; 12.1; 12.6(x4)
m.Yom. 6.1
m.Zeb. 2.3(x5); 6.7(x5); 12.5(x2)

Plural forms (26)
m.Yad. 4.7(x2)
m.Nid. 6.11(x2)
m.Ker. 1.1
m.Hul. 12.4
m.Men. 9.7; 3.7; 5.7
m.Hor. 1.1; 3.3(x2)
m.Ab. 2.1
m.Makk. 3.15
m.San. 8.1
m.Qid. 1.7(x2); 4.14
m.Ned. 2.2; 3.11(x2)
m.Meg. 1.9
m.Yom. 8.4
m.Bik. 4.2
m.Maʿaś. 5.11
m.Ber. 2.2

Syntagmatic analysis of the singular forms

Singular forms: 132 (CS 36; PS 32; AS 64)

1. Attributive function

Adjectives

ḥmwr 'heavy, important, stringent' (*Ab*. 2.1)
ql 'light, minor, easy' (*Ab*. 2.1; 4.2; *Hul*. 12.4)

Quantifiers

kl 'all, every' (*Qid*. 1.7x3; 1.9; *Makk*. 3.4; *Hor*. 2.6)
ʾḥt 'one, a' (*Ab*. 4.11; *Men*. 3.7; *Makk*. 3.15)

Pronominal suffixes

3rd singular masculine (*Bek*. 8.6; *Men*. 1.3x5;10.9x2; *Zeb*. 2.3x5; 6.7x5; *Meg*. 2.6x2; *Yom*. 6.1; *Pes*. 3.7)
3rd singular feminine (*Yeb*. 5.2)
3rd plural masculine (*Ber*. 1.1x3; *Ed*. 5.3; *Zeb*. 12.5x2; *Neg*. 14.4; *Par*. 12.4)

Reference: the offerings (*Ber*. 1.1x3; *Zeb*. 2.3x5; 6.7x5; *Zeb*. 12.5x2; *Men*. 1.3x5; 10.9x2); the faithful one (*Pes*. 3.7); the two goats of the Day of Kippur (*Yom*. 6.1); *birds* (*Neg*. 14.4); *daḇar* (*Meg*. 2.6x2); the ceremony of the *ḥaliṣà* (*Yeb*. 5.2); *mayim* (*Ed*. 5.3; *Par*. 12.4); the father (*Bek*. 8.6)

Nominal complements

Governing nouns or adjectives

ʾisur 'band, prohibition'
ʾswr mṣwh 'a prohibition on account of a commandment' (*Yeb*. 2.3x2; 3.2; 3.4)

bəʿilà 'sexual intercourse'
bʿylt mṣwh 'the sexual intercourse of *miṣwa*' (*Nid*. 10.1)

hepṣed 'loss, damage' (antonym of *śakar* 'salary, reward')
hpsd mṣwh 'the loss in carrying out a religious duty' (*Ab*. 2.1)

zahir 'observant'
zhyr bmṣwh qlh 'meticulous in a small religious duty' (*Ab*. 2.1)
(zhyr) kmṣwh ḥmwrh 'meticulous in large religious duty' (*Ab*. 2.1)

milḥamà 'war'
bmlḥmt mṣwh 'a war subject to religious requirement' (*Sot*. 8.7x2)

mṭ 'dead one'
mt mṣwh 'a corpse of *miṣwà*' (*Naz*. 6.5; 7.1)

śaḵar 'salary, reward' (antonym of *hepṣed* 'loss, damage')
śśkr mṣwh 'reward of religious duty' (*Ab*. 4.2)

šəḇuaʿ 'oath'
bšbwʿ mṣwh 'oath involving religious duty' (*Shebu*. 3.6)

šəya(y)r 'remnant, relic'
šyry mṣwh 'he residue of the requirement' (*Men*. 9.7; *Neg*. 14.10)

šaluaḥ 'messenger, agent'
šlwḥy mṣwh 'agents engaged in a religious duty' (*Sukk*. 2.4)

taglaḥat 'hair cutting'
btygḥt mṣwh 'cutting the hair for a religious duty' (*Naz*. 6.5)

Governed nouns

ʾezuḇ
mṣwt ʾzwb 'the requirement of the hyssop' (*Neg*. 14.4; *Par*. 11.9x2)

ben 'son'
wmṣwwt bnw 'and the requirement of (redeeming) his son' (*Bek*. 8.6)

gəʾullà 'redemption'
mṣwwt gʾwlh (*Bek*. 1.7)

ḥaľiṣà 'rite of removing the shoe'
mṣwwt ḥlysh (*Bek*. 1.7x2; *Yeb*. 12.1; 12.6)

yom
mṣwt hywm 'the religious duty of the day' (*RH*. 3.3x2)

yəʿidà 'designation, appointment, testimony'
mṣwwt hyʿydh (*Bek*. 1.7)

yibun 'levirate mariage'
mṣwwt hyybwn (*Bek*. 1.7x2)

lulaḇ 'lulav'
mṣwt hlwlb (*Suk*. 4.4)

neḥěnaqim
mṣwt hnḥnqym 'religious requirement of the strangulation' (*San*. 7.1)

neḥěragim
mṣwt hnhrgym 'religious requirement of the decapitation' (*San*. 7.1)

nisqalim
mṣwt hnsqlym 'religious requirement of the stoning' (*San*. 7.1)

niśrap̄im
mṣwt̄ hnśrpym 'religious requirement of the burning' (*San*. 7.1)

nəṭilà
mṣwt nṭylh 'the requirement of removing [the sinew of the hip]' (*Hul*. 7.2)
pədiyyà / *pədayyà* 'redemption'
mṣwt pdyyh (*Bek*. 1.7x2)

qali
mṣwt qly 'requirement of roasted grain' (*Men*. 10.4)

ʿăraḇà
mṣwt ʿrbh 'the requirement of the willow (branch)' (*Suk*. 4.4)

ʿăriṗà 'breaking the neck'
mṣwt ʿryph (*Bek*. 1.7)

ʿeṣ ʾerez
mṣwt ʿṣ ʾrz 'the religious requirement concerning the cedarwood' (*Neg*. 14.4)

ʿăšè
mṣwt ʿśh 'positive commandment' (*Shebu*. 2.3x2; *Hor*. 2.4; *Qid*. 1.7x2)

sukkà
mṣwt swkh 'the religious requirement of dwelling in a sukkah' (*Suk*. 2.7)

ʿomer
mṣwt hʿwmr 'the requirement of the *ʿomer*' (*Men*. 10.2; 10.9)

śəṛeṗà
mṣwt śryph 'religious requirement of burning' (*San*. 7.1)

Governed Pph

With the preposition *b*
mṣwwh blʾtʿśh 'negative commandment' (*Hul*. 12.4; *Hor*. 2.4; *Makk*. 3.4)
mṣwt hywm bšwpr 'the *miṣwà* of the day applies to the *šofar*' (*RH*. 3.3)
šmṣwwt hywm bḥṣṣrwt 'the *miṣwà* of the day applies to the trumpets' (*RH*. 3.3)
mṣwh bgdwl 'the religious duty for the oldest [surviving brother]' (*Yeb*. 2.8; 4.5)
mṣwh bdyynym 'the religious duty for the judges' (*Yeb*. 12.6)
mṣwh btlmdym 'the religious duty for the disciples' (*Yeb*. 12.6)

With the preposition *byd*
mṣwh byd gwʾl 'a religious duty in the hand of the avenger of the blood' (*Makk*. 2.7)

With the preposition *ʿl*
šmṣwwtw ʿl ʾbyw 'the requirement of redeeming him [the father] falls upon his father' (*Bek*. 8.6)
wmṣwwt bnw ʿl yw 'the requirement of redeeming his son falls on him' (*Bek*. 8.6)
mṣwh ʿl kl hʿwmdym 'the duty of all bystanders' (*Yeb*. 12.6)
wʿlyk mshwh 'the religious duty is yours' (*BM*. 2.10)

With the preposition *mn*
mṣwh mn htwrh 'a religious duty enjoined by the *Ṭorà*' (*BM*. 2.10)

Relative clauses

With the verb *ktb*
Kl miṣwh šktwb btwrh 'each commandment which is written in the *Tọrà*' (*Hor*. 2.6)

2. Predicative function

Miṣwà as a subject
qdm 'to advance, precede, be first, take precedence' (*Bek*. 1.7x4)
grr 'to draw' (*Ab*. 4.2)
tlh qal passive 'to depend on' (*Qid*. 1.9)
nhg 'to apply' (*Qid*. 1.9)

Verbs governing miṣwà as a direct object

bṭl 0/2 'to abolish' (*Shebu*. 3.6; 3.8)
grr 'to draw' (*Ab*. 4.2)
qwm 0/2 'to fulfil, carry out, execute' (*Suk*. 2.7; *Shebu*. 3.6; *Men*. 10.4; *Hul*. 7.2)
ʿśh 'to do' (*Qid*. 1.9x2; *Makk*. 1.7x2; 3.15x2; *Ed*. 5.3; *Ab*. 4.11; *Par*. 12.4)

Verbs governing miṣwà as an argument or adjunct

kwn T/2 *lšm* 'to intend for the sake of' (*Bek*. 1.7)
ḥyym mšwm/mšm 'to be liable on grounds of/by virtue of' (*Betz*. 5.2x2)
qrb k 'to offer according to (its requirement)' (*Men*. 1.3x5; *Zeb*. 2.3x5; 6.7x5)
rwṣ b 'to run after' (*Ab*. 4.2)
qdm l 'to advance, precede, be first, take precedence' (*Bek*. 1.7x4)
ḥzr b 'to return to, to go on to do' (*Pes*. 3.7)
śrp N/1 k 'to burn according to (its requirement)' (*Zeb*. 12.5x2)

Syntagmatic analysis of the plural forms

Plural forms: 26 (CS 6; PS 1; AS 19)

1. Attributive function

Adjectives/participles

ʾămuṛot battọrà 'written in the Tọrà (*Hor*. 1.1; *Nid*. 6.11x2)
ḥămuṛot 'heavy, important, stringent' (*Hul*. 12.4)

Quantifiers

kl 'all, every' (*Maʿaś*. 5.11; *Bik*. 4.2; *Meg*. 1.9; *Ned*. 3.11x2; *Qid*. 1.7x2; *San*. 8.1; *Hor*. 1.1; 3.3; 9.7)
šlš 'three' (*Men*. 5.7)
ʾrbʿ 'four' (*Men*. 3.7)

Pronominal suffixes

1st singular (*Qid*. 4.14, quotation from Gen 26:5)

Reference: The pronoun refers to *YHWH*.

Nominal complements

Governing nouns or adjectives

ʾḥt 'one' (Hor. 3.3)

rgyl l 'be used to' (*Yom*. 8.4)

ʿwl 'yoke' (*Ber*. 2.2)

ḥyyb b 'liable to' (*Bik*. 4.2; *Nid*. 6.11x2)

ḥwmr ʿl 'ritual restriction' (*Hul*. 12.4)

Governed nouns

ʾb 'father' (*Qid*. 1.7)

bn 'son' (*Qid*. 1.7)

ʿăśẹ̀ 'positive commandments' (*Ker*. 1.1)

2. Predicative function

Verbs governing miṣwot as direct object

ʿśh 'to do' (*Ned*. 3.11)
šmr 'to keep' (*Qid*. 4.14).
rbh H/1 hiqṭil 'to make many' (*Makk*. 3.15)
ṭ ʿn 'to require' (*Men*. 5.7)

Verbs governing miṣwọt as an argument or adjunct

bwʾ ʿl 'to transgress' (*Meg*. 1.9; 9.7)
ʿbr ʿl 'to transgress' (*Ned*. 2.2)

Glottodidattica

Language Teaching

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: D. Cappelli (2021) Dyslexia in L2 learning: comparison between languages and linguistic anxiety. *Qulso* 7: pp. 265-288. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12011

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Dyslexia in L2 learning: comparison between languages and linguistic anxiety

Deborah Cappelli
Università degli Studi di Firenze (<deborah.cappelli@unifi.it>)

Abstract:

In contemporary society, multilingualism is an increasing reality, and the study of foreign languages has a primary role in the educational policies of many countries. Furthermore, in the last years there has been a particular interest in specific learning disorders, so multilingual competence in subjects with dyslexia is an extremely interesting aspect to investigate. After a short introduction on learning disorders, this article examines the definition of dyslexia, both acquired and developmental type, and analyzes the neurophysical, neuropsychological and cognitive matrix hypotheses on the origin of the disorder. Given that learning to read and write is a fundamental stage in a child's life, reading and writing skills are explored in this work, together with the related disorders. In addition, the most problematic topics in learning a foreign language for dyslexics have been identified as well as the different impact that the characteristics of a language may have on the expression of the disorder, carrying out a short comparison between Italian and some foreign languages studied in school. Looking back over the main theories that have been dealing with this subject, it has been argued about the barriers that learners with dyslexia may find in learning a foreign language.

Keywords: *Dyslexia, Linguistic Anxiety, Orthographic Opacity, Phonological Awareness, Second Language Acquisition*

1. Introduction

Specific Learning Disorders (SLD) is one of the most relevant aspects not only in the psycho-pedagogical field, but also in the medical-pediatric field and it has been subject of study since a long time. In Italy, research and theoretical publications have been active since the 1970's, but it is in the last decade of the XX century that this topic has had an exponential increase. Postgraduate courses or Masters have been established, continuous research and a massive publication have been developed. Furthermore, it is worth mentioning the work conducted by clinicians in the context of the

Consensus Conference, summarized in the *Raccomandazioni per la pratica clinica dei disturbi specifici di apprendimento* of 2007, which defined the diagnostic path of these disorders, collecting theoretical and operational definitions to take action (Chiappetta Cajola and Traversetti 2017).

In the academic field the Italian law, Legge n. 170/2010, should be mentioned, which identifies dyslexia, anorthography, dysgraphia and dyscalculia as Specific Learning Disorders and also assigns both to the national education system and universities the task of identifying the most appropriate didactic forms and evaluation methods in order to make students with SLD achieve educational success. The legislation was then specified thanks to *Le linee guida per il diritto allo studio degli alunni e degli studenti con disturbi specifici di apprendimento* from MIUR (2011). In recent years there has been a significant increase in diagnoses compared to the past. This does not mean that before there were fewer subjects with SLD, but those who suffered from it were often labeled as "lazy"; today however the disorder has been recognized and a path has been identified in order to follow and facilitate the dyslexic in learning.

In the North American context, the term *learning disability* identifies various issues concerning cognitive development and school learning, "non imputabili primariamente a fattori di handicap mentale grave e definibili in base al mancato raggiungimento di criteri attesi di apprendimento rispetto alle potenzialità generali del soggetto" (Cornoldi 2007: 10). In Italy, the term collects various meanings as it can refer to very different situations: in the most general sense it includes any condition that expresses "special needs", while more specifically this term can be used to refer to cases of "slight intellectual disability" or cases in which specific learning disorders are evident and are related to problems in learning the skills of reading, writing and calculating (Zanobini and Usai 2008). Specificity is a characteristic of SLD, since the disorder affects a specific skill domain in a significant but limited way, and it does not affect the general intellectual functioning.

1.1 Specific learning disorders

Specific learning disorders concern issues that emerge in fields related to school education, such as reading, writing and calculating. In 2007 with the official presentation of *Raccomandazioni per la pratica clinica dei disturbi specifici di apprendimento*, developed using the Consensus Conference method, a significant step forward was taken in defining these disorders more precisely, here below listed:

- Reading disorder. Generally, it is an evolutionary dyslexia, a specific disorder in the automation (speed) and in the correctness of the reading.
- Writing disorders. We can distinguish them in two types:
  - Anorthography: is a disorder in the correctness of writing, understood as a process of transcription between phonology and graphemic representation of the word, to be distinguished from morph-syntactic accuracy.
  - Dysgraphia: is defined as a specific disorder in the manual realization of graphemes and therefore in graphism, of which calligraphy represents an emblematic exemplification.
- Mathematical learning disorders. These are disorders related to calculation problems (dyscalculia), logical reasoning, problem solving, and spatial reasoning (geometry). Sometimes it affects in a non-differentiated way its various aspects, which involve very heterogeneous cognitive processes that hardly have something in common between them. The inclusion in a single category comes from the fact that many students manifest such forms of generalized discomfort.[1]

[1] In the DSM-IV and ICD-10 diagnostic manuals, all forms of mathematical learning were put together, while the latest version of the ICD limits the disorder of calculus (dyscalculia) to its specific aspects, excluding all those related to other problems.

The Consensus Conference also identified the typical features of SLD:

- Developmental feature of the disorder.
- Different expression of the disorder in the different stages of evolution of the skill examinee.
- Constant association with other disorders (comorbidity).
- Neurobiological character of procedural anomalies characterizing SLD; they interact in an active way with the environmental factors in determining the appearance of the disorder.
- The specific disorder must have a significant and negative impact on school adaptation and/or everyday activities.

Talking about comorbidity, it should be noted that, in general, the condition of dyslexia is often linked to that of anorthography and these, in turn, are associated with a condition of dyscalculia or other neuro-developmental disorders, such as a specific speech disorder, attention-hyperactivity deficit disorder, or behavior or mood disorder. Comorbidity may not always be present, so before taking a position regarding the common etiopathogenetic models among these disorders, it is necessary to understand why they can manifest both in a joint or isolated way (Cornoldi 2007). Currently the most prudent position seems to be the one adopted by Consensus, which consists in considering the presence of disorders as co-occurrences without taking position about the hypotheses of a common etiopathogenetic factor.

In *Le line guida per il diritto allo studio degli alunni e degli studenti con disturbi specifici di apprendimento* of MIUR (2011), there is a reference to the individualized and personalized teaching method, which involves the use of compensatory tools and exemption measures, with the aim to ensure that students with SLD can achieve educational success. "Individualized" didactics means the intervention calibrated on the individual, rather than on the entire class or on the small group, which becomes "personalized" when addressed to a particular learner. Among the compensatory measures, educational and technological tools that replace or facilitate the required performance in the skill deficit, we can indicate: speech synthesis, recorder, and video writing programs with spell checkers, calculator and other technologically less advanced tools such as tables, formularies and concept maps. The exemption measures are instead interventions that allow the student not to perform some tasks which, due to the disorder, are particularly difficult and do not improve learning. An exemption measure may consist, for example, in not having a student with dyslexia reading a long text, as the exercise, due to the disorder, does not improve his/her reading performance. Or it is possible to allow students with SLD to take advantage of more time for the performance of the tests, or to be able to perform the test on a content that is disciplinarily significant but reduced, since the disorder engages them for longer than their classmates in the decoding phase of the test items.

2. Developmental dyslexia and acquired dyslexia

It is necessary to distinguish between acquired dyslexia and developmental dyslexia. In the case of acquired dyslexia it happens that a subject able to read without any difficulty, begins to make mistakes or fails to recognize words. Usually these are difficulties due to some pathological event that has resulted in lesions in the cortical areas involved in the transcoding process. It occurs in an adult subject, but there may be cases of acquired dyslexia even in children (Stella 2004).

Developmental dyslexia (DD) instead, it is found in children from birth and manifests itself at the beginning of the learning process of reading. In case of acquired dyslexia, the damage

can have much more limited consequences than the congenital condition. Often the first one only affects reading or some of its aspects, such as the recognition of new words, while in case of developmental dyslexia, the disorder is more extensive and especially at the beginning may affect the whole writing system.

The possibility of re-education is also different: in the case of acquired dyslexia it is a matter of regaining a function that the subject already possessed, while in the case of DD the subject must acquire a function that does not yet possess, by using a neurobiological system that has peculiarities that affect learning. Therefore, in the second case, there is nothing to repair or replenish, but measures must be used to facilitate the acquisition of the transcoding process and its automation. So, it would be more correct to talk about techniques of assisted education, rather than re-education.

2.1 Developmental dyslexia

If we want to define developmental dyslexia based on the criteria of definition by exclusion that characterizes the official nosography, DD can be described as "un disturbo manifestato nell'apprendimento della lettura nonostante istruzione adeguata, in assenza di deficit intellettivi, neurologici o sensoriali e con adeguate condizioni socioculturali" (Sabbadini 1995: 411). If we want to search for the primary characteristics of DD, the task of defining it is more complicated, as it is considered a complex disorder attributable to very different causes. According to some authors it is a defect of a specific superior cortical function, according to others it is attributable to a language disorder sometimes manifested, sometimes very difficult to identify before schooling, or a deficit in phonological processes, or it also could be caused by a defect in automation processes.

Beyond the hypotheses on the nature of the disorder, there are commonly shared elements such as:

a) DD is a specific disorder as it occurs as an isolated disorder compared to other cognitive performances, which are generally good, and in subjects free from other deficits or inadequate socio-economic and relational conditions.
b) It occurs in subjects without clinically detectable brain injury. It is present from birth, even if the most evident manifestations may appear only with schooling.
c) The clinical definition of DD often includes also the writing disorders (dysgraphia and anorthography) and those of numerical code and calculation (dyscalculia; Sabbadini 1995).

Developmental dyslexia shows itself as a learning disorder that persists throughout the subject's lifetime, although it changes in various degrees of expressiveness based on the severity of the disorder, the cognitive characteristics of the subject, and the educational and relational opportunities that the subject receives. DD initially shows itself with a difficulty in learning the writing and only in a second moment it affects also the reading, but during development there is a more substantial reduction in writing problems than in reading problems. By analyzing the performance of dyslexics, some researchers have hypothesized that their reading difficulties are in fact the manifestation of a learning delay and that, therefore, are not attributable to a specific disorder. As a matter of fact, with time and exercise, there is a significant improvement in the speed and accuracy of decoding for most of dyslexics, this would demonstrate that there is a delay and not a deviation from the course of normal development. However, despite improvements, developmental dyslexics always maintain significant decoding difficulties even when adults.

A child with a delay repeats continuously his mistakes, very slowly tends to eliminate them and he/she is able to memorize the new acquisitions, while the child with specific reading disorder fails to include in his/her diagrams the result of learning. Often it also happens that a new acquisition occurs at the expense of an old one that seemed acquired; these effects are not typical of the delay, but rather of a specific disorder. The World Health Organization has identified five conditions that must exist in order to define a reading disorder as developmental dyslexia. These are very precise criteria which must all be fulfilled (Stella 2004):

1. The intellectual level of the subject with reading disorder must be within the normal range (I.Q. greater than or equal to 85).
2. The reading level must be significantly distant from the one of a child of the same age or same attended class. It should be less than the II standard deviation expected for the age or class attended.
3. The subject must not have neurological or sensory disorders that may justify reading difficulties as an indirect consequence.
4. The disorder must be persistent, despite adequate schooling and specific educational interventions.
5. The reading disorder must have consequences on schooling or in social activities where the use of reading and writing is required.

Since from the earliest descriptions of children with dyslexia, it was hypothesized that dyslexics formed a group with peculiar characteristics, not only cognitive but also neurological. In the last decades of the last century, we have tried to investigate if the brain of dyslexic has structural alterations, carrying on studies on patients with developmental dyslexia, both through traditional methods such as autopsy studies, and new methods for the study of brain morphology and the functioning of the Central Nervous System, such as dynamic neuroimages: PET (tomography and positron emission) or MRI (functional magnetic resonance). Before the introduction of these new instruments it was the American neurologist Galaburda (1989) who discovered, through autopsy, the existence of small alterations of the brain areas involved in language processing, in order to demonstrate that all these small changes in the neuronal substrate could be the cause of functional deficits such as dyslexia. Specifically they have been described as: "sia alterazioni del pattern di asimmetrie cerebrali di aree linguistiche, sia malformazioni corticali minori della corteccia perisilviana. In particolare, il planum temporale era simmetrico in tutti i cervelli esaminati, mentre nel cervello normale presenta nel 70% dei casi un'estensione maggiore a sinistra che a destra" (Sabbadini 1995: 415).

The morphological alterations of the perisylvian cortex highlighted by the works of Geschwind and Galaburda (1985) concerned the presence of ectopias, a term used to indicate that neurons are not in their place. In fetal brain, during pregnancy, it happens that cortical neurons move in a wide migratory movement, from the area where they are formed, up to their final position in the different layers of the cortex; in dyslexic neuronal migration is abnormal (Dehaene 2007).

2.2 Hypotheses on dyslexia

The theory of neurodiversity is based on three fundamental assumptions. The first is related to the fact that the brain must be understood as an ecosystem in which the individual parts (such as the ability to organize, to face the problems and to adapt) are compensated and integrated one

with each other. The second assumption is that the disability model is not universal; as a matter of fact, dyslexia is seen as a deficit in cultures that center on literacy by considering writing and reading as predominant abilities. The third assumption, finally, states that success occurs when the brain and the environment adapt to each other; the brain adapts to the requests of the environment, the environment to those of the brain. Failure leads to mismatching and social disadvantage, unrelated conditions from the neurological disorder, but which have much to deal with the will and the ability of the people around the subject to welcome and enhance this dysfunction.

The first hypothesis is neurophysical, whose studies date back to the 1960's but arrive up to the present day; as a matter of fact, in the last fifty years, brain regions that could form the basis for the neuroanatomical origins of dyslexia have been identified. This theory analyzes the individual areas of the brain and the mechanisms that could lead to dyslexia, highlighting particularities in the following areas:

- Cerebral cortex: in a normally developed brain the left hemisphere is larger than the right one, while in the dyslexic subject this asymmetry is not detected, which could be characterized in a perfect symmetry or again in an asymmetry in reverse (Geschwind and Galaburda 1985). In addition, the brain of a dyslexic person shows very small differences in the location of brain areas, and these small differences are more marked in the area intended for language (Galaburda 1989, 2005).
- Magnocells: other studies have reported signs that the dysfunction may result from a small difference of the "linguistic signal" transmission. Both auditory and visual signals are transmitted to the brain in two different ways; through the magnocellular path, that is where the signal is transmitted quickly, and the parvocellular one, that is the most important one for the elaboration of the details. In the dyslexic subject the difficulties are at magnocellular level, consequently the information is not transmitted at the necessary speed (Talcott *et al.* 1998, 2000; Stein 2001, 2003).
- Cerebellum: cerebellum is a part of the brain considered to be the one that allows the functioning of the set of abilities, both motor and linguistic-cognitive. With regard to language, the cerebellum seems to be involved at least in oral production and reading. This would also explain why children suffering from this dysfunction report differences in eye movement during reading, and some awkwardness in coordination and fine movements (Nicolson and Fawcett 2008).

Another hypothesis is the one of neuropsychological origin, which has been involved in analyzing the different types of dyslexia and has led to identify three types of dyslexia:

- phonologic, linked to grapheme-phoneme conversion. This happens when the graphic form of the words has not yet been memorized by the student or if the student encounters unknown words. The consequences of this type of dyslexia are a low degree of phonological awareness and difficulty in deciphering new words;[2]
- superficial, that is the ability to connect the graphic form, the phonological one and the meaning of an item. The main result of this dyslexia is the slow reading;
- deep dyslexia, which includes both ways. This type is not very frequent and can also cause semantic errors.

[2] Phonological awareness refers to the ability to manipulate the sounds of language through processes of fusion, deletion and synthesis of phonemes.

This categorization has received confirmation regarding of the acquired dyslexia while, in order to frame also the developmental dyslexia in this descriptive model, it is essential to consider the following variables:

- dimension of development: learning has always taken place under altered conditions, therefore in the evolution of differentiation the subject can use different strategies according to situations and needs;
- formation of the cognitive and learning style: each subject develops his own strategies to solve the problems of everyday life, starting from his/her own abilities and avoiding his/her own limits. The dyslexic subject develops these skills by maturing a style that best suits him, whether it is a visual-semantic or auditory-phonological one (Daloiso 2012).

Finally, we find the hypothesis of a cognitive matrix that attributes dyslexia to individual differences in phonological processing (Snowling 1987). At the basis of a correct development in the field of reading and writing, a previous development of phonological awareness is necessary, which allows a set of operations necessary for written encoding and decoding. Numerous studies confirm the link between dyslexia and low phonological awareness, but this hypothesis is not able to explain all the difficulties of the dyslexic subject. Above all, it cannot explain all the differences, for example why a child reads slowly but accurately and another child reads the exact opposite. In this regard, the hypothesis of the double deficit tries to provide an answer by stating that at the origins of dyslexia we can find two very distinct deficits: the first one concerning the phonological side, the other one concerning the rapidity of recovery of already known words. The first problem leads to inaccuracy in reading the second instead to a slow but accurate reading (Wolf and Bowers 1999).

Among the most recent research there is *the hypothesis of the automation deficit* (Nicolson and Fawcett 2008) which resumes the weaknesses of other theories, especially those related to the phonological deficit and double deficit. Both theories, as a matter of fact, focus on the linguistic dimension and do not explain the difficulties that a dyslexic subject may encounter in the non-linguistic tasks, for example: calculate quickly, remain in balance on one foot while being blindfolded, etc. This theory consists in considering dyslexia as a lack of automation of some processes; the listed tasks have some common denominators with normal dyslexia disorders:

- the activities must be automated, so that they can be done without thinking too much about what should be said or done;
- involve implicit memory,[3] which allows the unconscious internalization of some procedures, which, once learned, require a minimum effort (Cardona 2010);
- involve the cerebellum.

Learning a foreign language (L2) involves interfacing with different cultural models, so that different cognitive scripts[4] may correspond to the same situation in a different culture. The main difficulty for foreign language students is to transfer the scripts acquired in the native language into the new culture they are studying with the further complication of cultural plurality, as in the case of English and Spanish and, in a less marked way, in German. By implementing

[3] Implicit memory allows the internalization of procedures that mind and body have learned to implement in an increasingly automatic and unconscious form and are stored there much of the oral skills in the native language: both idiomatic expressions, both words and recurring communicative acts are memorized unconsciously.

[4] Theory of cognitive scripts (Nelson 1986), according to which, interacting with the environment, the child learns to associate its pragmalinguistics rules to each communicative situation.

the vision of Nicolson and Fawcett (2008), dyslexic student may not have fully automated the scripts of his/her own culture and learning new *scripts* would be an advantage allowing him/her greater adaptability and flexibility.

Building a picture of all three hypotheses together is quite complicated, since the research did not provide clear conclusions yet, but it has highlighted aspects that can be shared such as: the fact that dyslexia leads to the formation of a learning style; that it originates from a neurodiversity and causes a difficulty in the automation of some tasks, the most difficult to automate are the ones of linguistic nature (phonological and orthographic dimension), mnemonic nature (counting and naming things quickly) or mobility nature (coordination or complex tasks).

3. Analysis of reading and writing skills

The learning of reading and writing is a fundamental step in a child's life, who is surrounded by the written language since childhood, so learning to decode writing is essential to interact with the environment. These skills combine written signs with meanings by exploiting the linguistic knowledge of the speaker, specifically the phonological, morphosyntactic and lexical knowledge that is part of his/her mental grammar or internal language. It is assumed that learning, reading and writing skills have a separate status from the acquiring of language system. While the acquisition of native language depends on an innate, biologically determined system, reading and writing are the result of a teaching *ad hoc* that involves the child's general learning abilities (Baldi and Savoia 2017). In this sense, writing and reading are considered as a product of the cultural evolution of man:

> Non si tratta, perciò, di meccanismi cognitivi innati per la lettura/scrittura ma il risultato dell'utilizzo di capacità cognitive preesistenti, specializzate per il riconoscimento di oggetti. È con l'addestramento che gli esseri umani hanno imparato a sfruttare queste capacità per riconoscere le forme delle lettere e delle parole e tradurre i suoni nel linguaggio scritto. (*Ibidem*: 110-111)

The alphabetical system is used to write European languages and nowadays many other languages of the world. It comes from Roman writing and is based on a conventional relationship whereby each letter or certain combinations of letters corresponding to a phonological segment. It should also be added that alphabetical writing systems may have irregularities, as they have been established on the basis of conventions which have changed over time. Even in a language with transparent and regular writing as Italian there are discrepancies in the correspondence between letters and sounds (Baldi and Savoia 2018).

Learning to read consists in connecting two brain systems present in the child: the visual perception of forms and the language areas. According to the model of neuronal recycling, writing is progressively fixed in the brain of the future reader and it must be placed in an optimal space where already functioning circuits are reconverted to another function. Subsequently, there is a reconversion which involves the transformation of a function which had its own usefulness in our evolutionary past, into a new and more useful function in the present cultural context (Dehaene 2007). In the specific case of reading learning, it takes place a reconversion of the networks of neurons initially dedicated to the visual perception of objects; Reading should progressively converge towards the left occipitotemporal region, and later take, over the months, a specialization for the written text and the interconnection with the other temporal, parietal and frontal regions.

Dehaene (*ibidem*) also analyzes the period preceding reading learning, as he argues that learning to read is possible only because the child's brain already has acquired a large part of

the appropriate neuronal structures. He also adds that the linguistic and visual development of the child, even before he begins to read, plays a fundamental role in the good preparation of the brain for reading. Since the first months of his/her life, he/she possesses an exceptional competence for the analysis of language and it is shown that even from the first day of life, he perceives linguistic contrasts and has a very special attention for the prosody of his/her native language. In addition, brain imaging of 2- or 3-month-old children has shown that the child's language skills are already resting on a cortical network of the left hemisphere, the same that is activated in the adult during language processing. Specifically, in our childish brain, it happens that:

> The upper left temporal region analyzes the sounds of the language, while the upper left temporal sulcus already shows a hierarchical organization, probably linked to a progressive analysis of phonemes, words and phrases. Even the lower left frontal region, the Broca area, a region traditionally implicated in the production of the word and in the analysis of grammar, is activated in the three-month-old child during the listening of sentences, even if at that age he does not produce anything but indistinct sounds. (*Ibidem*: 229)

These initial preconditions are subsequently transformed into learning; During the first year, the network of language areas gradually specializes under the influence of the native language. As soon as the child begins to read, in the age of 5 or 6, he/she has a detailed representation of the phonology of his/her language, a vocabulary of a few thousand words, and he also knows the main grammatical structures. Obviously, he does not know that he has the rules of representation, but they are already in the set of neuronal circuits of oral language, ready to deal with the written text.

At the same time, the child's visual system is structured. We know that since the first days of life, he/she pays particular attention to faces, which seem to activate, as in the adult, a part of the right occipitotemporal region. This specialization becomes even greater during the first year of life and continues for about ten years. Around the age of 5-6, when the child learns to read, it is likely that the ventral visual system is still in a period of intense plasticity in which the functional specialization is far from being stabilized. This is a particularly favorable phase for the acquisition of new visual objects such as letters and words.

To learn how to understand and speak their native language, children do not need any particular teaching, although adults perform an important tutoring function that, if properly exposed to the language, children will be able to speak it sooner or later. The written language, on the other hand, is based on a complex system of correspondences between sounds and symbols. Therefore its learning is not natural but needs a direct teaching that explains the rules necessary to interpret the graphic symbols and bring them back to the knowledge that the child already owns.

Understanding a word in oral form is very different from decoding a written word, however both processes require the activation of lexical foreknowledge. As a matter of fact, every time we listen to or read a word our memory is activated, which gives us access to our lexical foreknowledge and allows us to understand the meaning of what we read or hear. Although the studies in the psychology of reading have not produced numerous developmental models dedicated to the description of the stages of development crossed by the child in his/her path as a reader, the main evolutionary theories allow us to outline the essential stages of reading-writing, composed by distinct but connected abilities:

1. Logographic phase.
2. Semi-phonetic phase.
3. Alphabetical phase.

4. Spelling stage.
5. Lexical phase.

This path is characterized both by *bottom-up* processes, according to which the child proceeds from the formal analysis of the letters to their semantic interpretation, and by *top-down* strategies, in which the child starts from semantic interpretation in order to arrive to the analysis of form and they are particularly productive. The context is, in fact, an essential variable right in the early stages of reading learning: the inexperienced reader seems to use more contextual information to decode the text (Camaioni 2001). With the lexical phase the reader can be considered able to process the written code automatically because, thanks to the construction of a large lexical background, the child will achieve an accurate reading-writing skill without the need to resort to the grapheme-phoneme conversion strategy if not for unknown words.

In accordance with the adult[5] models, we can assume that upon completion of this evolutionary path, a specific module is formed into the child's mind in order to manage decoding of writing. The modular perspective to cognitive processes, inaugurated by Fodor (1983), postulates the existence in the human mind of distinct and independent modules that manage the higher cognitive functions such as attention, vision, language, etc. The theory of modularity has an innate nature as it states that such modules are innate, that means they are inscribed already in the genetic patrimony of each person. However, later it was discussed that some of the modules could be acquired on the basis of a particular external input, as in the case of reading. This hypothesis then had numerous confirmations in the study of subjects with acquired dyslexia, who presented particularly specific deficits in some reading processes and not in the other skills.

This module would work through two distinct and autonomous procedures:

a) The visual path: through which it is enough to recognize letters to recover the orthographic representation of a word as well as the meaning and phonetic form. It can only be used for known words, whose spelling, semantics and phonetic information are already stored;
b) The phonological path: the subject recognizes the letters, recalls to the memory the associated phonological units and reconstructs the correct pronunciation of the word. The expert reader only uses it in the presence of unknown words.

Being able to decode a written text is a necessary but not sufficient condition to understand the content; *bottom-up* strategies guarantee superficial decoding (from detail to general), while *top-down* strategies (from general to particular) allow the understanding of content. However, it is necessary to emphasize that the two ways are not sequential, in fact it is not true that the text is first decoded and then it is understood; often the two processes are simultaneous and influence each other (Daloiso 2012).

3.1 Reading disorder

As we have already seen, reading-writing disorders are one of the largest categories of learning disorders, (cfr. Sections 2.1, 2.2); the first issues to be highlighted regarding the pro-

[5] In particular, the modular model of double access postulated by Coltheart (1981), which is still the most accredited in the neuropsychological field.

cess of reading and understanding a text are two: the first is one can be linked to a difficulty in the decoding of written words, that is the ability to read a text fluently and correctly; while the second one consists in the understanding of what is read, that means a set of skills that allow the subject to create an image of what he is reading. In this case it should be noted that a child with reading disorders will have no problem in understanding a written text if it will be presented to him in an oral way.

Children with problems understanding a written text require specific attention in order to develop that ability, and have a lower performance than children without this dysfunction, but it is still possible to spot some difficulties:

- Linguistics: as, for example, in case of an unknown word it is possible that the student may not be able to analyze it on the morphosyntactic level, by assigning it to a category (nouns, adjectives, verbs, etc.). He may also have difficulty in putting it into the right context and thereby demonstrating the understanding dysfunction;
- Textual: reading a text and searching for certain answers and information;
- Mnemonic: that is to be able to capture the important information and to omit the irrelevant ones: thus creating an overload of information and having even more difficulties in the study;
- Metacognitive: reading requires the activation of strategies that allow the student to control in a conscious way the process of understanding, something in which a dyslexic subject is lacking.

Another problem of reading is that of "letter by letter" dyslexia, in other words, when a child presents this disorder, he/she can read a word only after having named the individual letters that compose it. The student must be able to perform visual analysis of the individual grapheme without being able to reproduce the sound of it and link it with the other letters.

Many disorders generally associated with written language, such as dyslexia, also affect oral skills, are related to linguistic input and should not be confused with cognitive problems. Understanding the language input is crucial during the first phase of learning grammar, when a consideration on the language is required. It is exactly in this phase the function that Noam Chomsky calls Language Acquisition Device (LAD) is activated, adapted to the glottodidactic practice by Krashen and Terrel (1983). In the oral comprehension of a linguistic input, for some students with dyslexia it is difficult to segment the units of meaning in the communicative flow and to reach a real metalinguistic competence. These difficulties in oral comprehension will be present in the student throughout school, with particular incidence at the stage of first exposure to languages, a period in which basic language competence is formed. In addition, the student perceives a higher speed of speech than the rest of their classmates and over the time it becomes a disadvantage that will generate frustration, low self-confidence and decreased motivation (Triolo 2016).

The useful element to predict the learning success of a L2 are: age, quantity and quality of language exposure; in cases of inadequate development of the L1 it is necessary to operate a hyper-learning of the L2 in order to provide the necessary linguistic material to reach the minimum level identified by Cummins[6] (Celentin 2016). Generally, the learning of the L2 takes place in

[6] There is a minimum level of fluency that must be reached both in L1 and in L2 to allow bilinguals to make the best use of cognitive and linguistic stimuli. The development of both languages allows growth from the linguistic and cognitive point of view, suggesting the hypothesis of inter-dependence between languages affirming that what

formal contexts in which, as we have already pointed out, the teacher plays an important role, so his/her choices at the methodological and didactic level have a big influence in the learners with dyslexia. Therefore, it is essential to understand the individual differences and considerate them in educational choices in order to avoid the creation of additional barriers for the learner. There are three different approaches that gave results in the last seventy years for dyslexic learners: the structuralism approach, the communicative approach and the educational-communicative approach.

4. Dyslexia in comparison between languages

Most studies on dyslexia are Anglo-Saxon, therefore, the prospect of difficulties was initially believed as universal, without considering the effect that phonological properties, orthography and morphology of a specific language may have in the development of reading skills. Thanks to the cross-linguistic data that we have, we know that some languages actually amplify the manifestations of dyslexia, while others attenuate them. The subject's skills may be more or less developed, depending both on the foreign language he decides to study and his native language. As a matter of fact, if a child chooses to learn a language very similar to his/her own, he/she will have less difficulty because he has already acquired the necessary strategies in his/her own language. However, if the chosen language is very different from his/her own the issues will increase, obviously. The factors affecting the relationship between dyslexia and foreign language are the following (Daloiso 2012):

- Orthographic transparency: affects the manifestations of dyslexia, so theoretically, a dyslexic student of a transparent orthographic language, while studying a shallow or deep-orthographic language can show signs of slowness but also of non-accuracy, a characteristic that generally he/she does not have in his/her native language;
- Linguistic affinity: which means the degree of similarity between two languages, which usually depends on a genetic relationship or typological similarities, and can be:
  - phonological: the degree of similarity that can be observed by comparing the phonological system of the native language to the foreign one;
  - morphosyntactic: for morphosyntax we mean that set of properties related to the nominal and verbal systems of a language that makes it similar or not to the student's native language;
  - lexical: that means the degree of formal and semantic similarity between the words of a foreign language compared to the native language.

The case of English shows that there is a certain degree of lexical affinity with Italian due to the many words in the English lexicon that have origins from French, Latin, Greek and Italian; in addition, the Italian language is embracing in its lexical repertoire many terms from English, also known to those who do not know the language. Italian has strong phonological, morphosyntactic and lexical affinities with both Spanish and French; on the contrary, it has as a very low affinity with German.[7] The fact that some languages amplify or mitigate the manifestations of dyslexia is linked to the fact that some languages encourage transfer[8] of those

is learned through one language is transferable to the other. Only the superficial aspects of languages are different while the aspects related to cognitive development are the same.

[7] For example, German and Italian are two languages distant from the morphosyntactic point of view because German has rules of lexical composition that needs the use of declinations.

[8] The linguistic transfer consists in moving the habits that have been consolidated in their native language in L2.

strategies already acquired in L1, while others force the student to learn other strategies more functional to the new language. In *Linee guida per il diritto allo studio degli alunni e degli studenti con disturbi specifici di apprendimento* by MIUR (2011) is reported that:

> Poiché la trasparenza linguistica […] influisce sul livello di difficoltà di apprendimento della lingua da parte degli studenti con DSA, è opportuno che la scuola, in sede di orientamento o al momento di individuare quale lingua straniera privilegiare, informi la famiglia sull'opportunità di scegliere – ove possibile – una lingua che ha una trasparenza linguistica maggiore. (19)

However, these considerations should not lead to the wrong conclusion that some languages are too difficult to be learned; with a good teaching methodology it is possible to learn them in the proper way (Balboni 2008). In addition, it is possible to make a comparison between Italian and foreign languages studied at school with the aim of creating a scheme to guide teachers in understanding the degree of difficulty that may involve studying the foreign language for the student.

	Trasparenza ortografica	AFFINITÀ LINGUISTICA CON L'ITALIANO		
		Fonologia	Morfo-sintassi	Lessico
Francese	Moderata	Moderata	Alta	Alta
Inglese	Bassa	Bassa	Moderata	Moderata
Spagnolo	Alta	Alta	Alta	Alta
Tedesco	Alta	Bassa	Bassa	Bassa

Table 1. Orthographic transparency and linguistic affinity with Italian (Daloiso 2012: 75)

While reading patterns in L1 allow us to formulate hypotheses and to explain the relationship between low levels of comprehension and existing cognitive-linguistic processes, the same cannot be said for research related to the comprehension of the written text in L2. If the process of reading worked in a similar way, the models analyzing the difficulties could also be applied to multilingual students. Instead, differences between languages can cause a different development in skills involved in the comprehension of a written text, depending on the linguistic background. Research aimed at identifying similarities and differences in different linguistic contexts is crucial in determining whether predictive models related to L1 can also be applied to multilingual contexts and whether they need to be modified or integrated (Grech, Sadeghi and Everatt 2016).

Starting from an analysis of the most relevant linguistic theories regarding the universality of reading processes, we can say that the written decoding is based on a universal processe that presents elements of variability linkable to the areas of phonology and spelling, which depend on the target language that you want to learn. The potential impact of these variables is particularly significant in the study of learning how to read in a foreign language and deserves a careful consideration.

4.1 Phonological awareness

Phonological awareness is one of the central aspects in most disorders and to be able to make a diagnosis it is necessary to know the age of the first exposure to the L2. As for phonological

variables, most psycho-linguistic studies say that the stages of development of phonological awareness are more or less the same in all children, despite the language they are exposed to. Therefore, the path proceeds from the highest level, from the word to the syllable, up to the analysis of individual phonemes. The *Grain Size Theory* (Ziegler and Goswami 2005)[9] stands out in the field of psycholinguistics studies and is based on the observation that the human brain is not programmed to read but rather to understand the communicative messages transmitted in the oral way. The brain, therefore, is not only able to process a language at a phonological level from the early stages of development but also uses phonological processing to learn how to decode the script. The human being has developed an area of the brain engaged in visual analysis of the form of words, which seems to be connected with the area of phonological analysis (Paulesu *et al.* 2010).

It is known that syllabic awareness develops before the learning to write begins (Ziegler and Goswami 2005), however, different languages can make children more sensitive to different syllabic structures: for example, while in many languages including Italian, Spanish and Chinese, the most frequent syllabic structure is CV, in other languages such as English and German more articulated structures are common. As for English, the most frequent syllabic structure is CVC (dog), which consists in over 40% of monosyllabic words, followed by CCVC (slow), CVCC (fast) and CCVCC (climb). This shows that each child identifies the most frequent syllabic structures in his/her language in a very simple way (Daloiso 2012).

The ability to analyze syllables is a universal stage, but the speed of development depends on the type of language. Again, the comparison between Italian and English is emblematic: on the one hand, Italian is an isosyllabic language, so syllables are to be pronounced with the same duration; on the other hand, English is a stressed-time language which means that the stressed syllables have a longer duration, while the others are reduced. Therefore, we can say that, unlike in Italian, in English "mission", "omission" and "emission" if pronounced in a context are practically homophonic, because the unstressed syllables have a reduction of the vowel, which is pronounced with an indistinct sound. In a stressed-time language, such as English (or German), the child has more issues in distinguishing the different syllable in order to use it as a unit of phonological analysis: it follows that syllable awareness emerges earlier in children of isosyllabic languages than in the ones, of the same age, of stressed-time languages (Duncan *et al.* 2006).

Another aspect that affects the development of phonological awareness is the sound profile. Each language is composed by a large set of sounds: the vowels are the loudest sounds that we can produce, followed by the semivowels, for example /w/ that we find in the Italian *uomo* and in the English "wine". Liquid and nasal consonants are more sonorous (/l/ /r/ /n/ /m/) than the deaf obstructing ones like /p/. If we analyze the composition of these sounds in different languages, it appears that Italian is a very sonorous language, considering the high distribution of vowels both at word level, considering that almost all Italian words end in vowel, and at syllable level being CV the most frequent structure. English is, in some ways, an opposite case to Italian, considering the predominance of consonant sounds at the end of the syllable and word.

The sound profile of each language affects in two ways the development of children's phonetic discrimination skills: the first influence has to take into consideration each specific language, meaning that the child, exposed to a more sonorous language, is more able to dis-

[9] Zieger and Goswami's theory believes that to understand cross-linguistic differences in reading, it is necessary to go into the meta-phonological development of different languages, at various levels (grain sizes) starting from the syllable up to the single phoneme, which are the universal characteristics of all languages.

tinguish the most sonorous syllabic structures. A second influence can be considered universal and could be resumed as follows: in any language, within the same word, the child manages to differentiate in a better way the syllables composed by contrasting phenomena at the sound level than the ones with many sound phonemes inside. For this reason, a child can perform more easily the phonological analysis of the word "cat" than "while", because in the second one the phonemes are all very sonorous.

A third variable that seems to influence the level of phonological awareness in children is the phonological closeness, or degree of similarity, between two words: apparently children develop a greater capacity of discrimination for groups of words characterized by a high phonological proximity, for example in Italian *casa, caso, cosa, cola, calo*; in English "sin", "keen", "mean", "bean". This is due to the fact that in each language some combinations of sounds are more productive than others, so the child must sharpen his/her discriminatory skills precisely towards those more frequent sounds that generate distinctions of meaning. It is also noted that in many cases phonological proximity produces rhymes, so exposure to specific combinations of sounds in a language also influences the ability to discriminate and create rhymes.

These data reveal that cross-linguistic variability in the development of phonological awareness affects more the contents in terms of sounds, which the child is able to manipulate, depending on the phonological properties of each language. On the other hand, there are no significant differences in the type of phonological analysis skills, since all children, at about the same age, are able to manipulate the sounds of their languages at the level of syllable, incipit-rhyme and phoneme. A substantial cross-linguistic difference emerges, however, at the level of manipulation of individual phonemes. While the ability to analyze the syllable and the incipit-rhyme structure emerges from the age of four, the ability to analyze individual phonemes seems to emerge at different times, in different languages, and appears influenced by contact with the written code. As a matter of fact, since in natural communication we do not hear single phonemes but a continuous flow of sounds, children are able to isolate single sounds when they start to relate them to the corresponding letters. It has been noted that in languages with greater coherence~~,~~ in the connections between letters and sounds, the ability to manipulate individual phonemes appears earlier, already at the turn of the fifth and sixth years of life.

An interesting aspect is that in multilingual subjects with linguistic disorders, the areas of fragility identified in one language are also present in the other acquired language. For example, a subject with a phonological awareness deficit in L1 shows elements of fragility in the same area even in L2, though the manifestations of fragility may change depending on the characteristics of L1 and L2. In fact, even in the case of similar languages, such as Italian and Spanish, slightly different manifestations have been discovered for the same disorder.[10] It has been suggested that the non-coincident linguistic behaviors of Italian and Spanish subjects with a linguistic disorder are caused in part by the different prosodic properties of the two languages[11] (Celentin and Daloiso 2017).

4.2 Orthographic transparency

From the point of view of the relationship between pronunciation and spelling, it is possible to distinguish between transparent orthography languages, or superficial languages, which

[10] As for morphosyntax, for example, Spanish speaking children tend to replace clitics and articles with other inappropriate forms, while Italian children of the same age tend to omit these forms.

[11] In Italian, the distinction between strong and weak syllables is not very marked as it is in Spanish.

tend to be uniquely related to phonemes and graphemes, and shallow or deep-orthographic languages, which are characterized by an even more complex relationship between oral and written form. To the first category belong Italian, Spanish, German, and Greek, and to the second group belong French and English. This is summarized in the following classification table of alphabetic languages according to the opacity of their spelling system and syllabic structure (Jiménez 2012: 27).

		Orthographic Opacity				
		Transparency				Opacity
Sillabic Structure	Easy	Finnish	Greek	Portoguese	French	
			Italian			
			Spanish			
	Complex		German	Duch	Danish	English
			Norwegian	Swedish		
			Icelandic			

Table 2. Classification table of alphabetic languages

The orthographic depth hypothesis, which emerged in the 1990s, initially argued that the degree of complexity of the relations between phonemes and graphemes in a language influenced the development of reading strategies: using the double-entry model of Coulthard (1981), transparent orthography languages would support sub-lexical strategies, that are phonological, thanks to the high correspondence between pronunciation and writing, while shallow or deep-orthographies languages would support lexical strategies, or the reading of the words "in block", since the phoneme-grapheme conversion does not work because of the irregularities of the language.

However, the original formulation of this theory did not find any solid experimental evidence: in particular the "strong" hypothesis does not explain why in transparent orthography languages, as in Italian, children are able to understand the correct accent of words, which is often not graphically marked: for example *leggere* can be both a verb or an adjective and only reading the words "in block" can clarify this ambiguity.

In its most recent formation, the orthographic depth hypothesis argues that in all languages, shallow or deep-orthographic ones or transparent orthography ones, reading involves both lexical and sub-lexical processes; however, the degree of activation of these strategies changes according to the type of spelling: transparent orthography language readers, as a matter of fact, seem to use more the sub-lexical (phonological) method, while shallow or deep-orthographic language readers of the same age, primarily use the lexical (visual) method. This "weak" version of the hypothesis seems to find confirmation in several languages, including Italian, English, German, Spanish, Greek and Turkish (Daloiso 2012).

In addition of affecting the types of processes, the degree of orthographic transparency also influences the speed of development, that is the time that children spend in the different stages of the learning path previously listed. This means that a 7-year-old Italian student generally reads with an almost perfect degree of accuracy, while an English child reaches the same level around the age of 10-11. Furthermore, according to a comparative research aimed to investigate the degree of reading accuracy in a sample of children from 14 European countries at the end of

the first year of literacy, German subjects were found to be 98% accurate, Italians and Spanish were 95% accurate, and the British were in last position with 34% (Seymour, Aro and Erskine 2003). So, although the specific reading disorder has a neurobiological origin, linguistic and environmental variants affect its characterization and frequency. A study showed that 47% of dyslexic Spanish speakers showed superficial dyslexia and only 22.8% phonological dyslexia, while the opposite occurs with English speakers with dyslexia. In addition, it has been found that languages with a transparent orthographic system have a greater decrease in reading speed and a less difficulty in accuracy (Matute and Guajardo 2012).

Cross-linguistic research suggests to consider the orthographic depth as a variable (Frost and Katz 1992), alternatively, to consider the degree of complexity of the relations between phonemes and graphemes in the spelling of a language. English, with its 44 phonemes and hundreds of possible spellings, is a deep-orthographic language, opposite to other languages such as Italian and Spanish; for example, the Italian phonological system has about 30 phonemes, corresponding to 21 graphemes. Italian has a high degree of orthographic transparency, which facilitates the conversion grapheme-phoneme in reading, despite the correspondence between symbols and sounds is not always perfect. This aspect simplifies learners of the same degree of transparency languages who can activate already acquired mechanisms, but it also facilitates learning for all the foreign learners who, due to their limited lexical background, choose the phonological pathway to decode. Italian is an alphabetical language with progression from left to right, which facilitates students with European native language, but can create representation problems for Arabic speakers, which have the opposite progression, for Chinese speakers, due to logographic writing, and for the Japanese ones, whose mixed alphabetic writing is ideogrammatic and with progression from top to bottom. Additionally, capital letters in Italian are distinctive and identify proper names and the beginning of sentences, but in some European languages their use is different, for example in German are used for all nouns, while other languages such as Arabic do not distinguish between upper and lower case letters, therefore, problems may arise in the use of punctuation (Melero Rodríguez 2015).

Spanish is also considered one of the languages with a transparent orthographic system for reading, which consists of: 27 graphemes, 5 digrams and 2 diacritic signs (acute accent and dieresis). This helps reading learning especially for Spanish children, who at the end of the first year of primary school correctly read 94.72% of words and 88% of pseudo-words. In contrast, English children read only 33.89% of words and 29.26% of pseudo-words correctly, due to the deep-orthographic system (Matute and Guajardo 2012). Already in 2002, the study conducted by Defior, Matos and Cary showed that the reading time of pseudo-words by first and second grade Spanish-speaking children is lower than the reading time of Portuguese children, whose language, even if it has a transparent orthographic system, is deeper than Spanish. However, it is important to add that the transparency of the Spanish orthographic system concerns reading but not writing. The phoneme-grapheme conversion is less consistent because the phoneme itself can be represented by two or more graphemes without any rule determining which one is the most appropriate. As a matter of fact, Spanish-speaking children acquire reading skills before writing skills and achieve the roof-effect in reading as early as the second year of primary education, while for writing we must wait for the fourth grade (Arfè, Dockrell and Berninger 2014).

French is considered as a medium opaque language with 190 graphemes in order to make 35 phonemes; for example, to write the sound [o] in French we can find the following twelve graphic combinations: *o*, *ot*, *ots*, *ocs*, *au*, *aux*, *aud*, *auds*, *eau*, *eaux*, *ho*, ô. French has conversion rules which, once they have been automated, make it possible to read and write even unknown words correctly. Nevertheless, for a dyslexic student French is particularly difficult because through these phoneme-grapheme conversion rules it is possible to write correctly

only about 50% of the words. The remaining words can only be read correctly through the lexical way, that is, having already in mind the sound image of the word itself. These are words defined as irregular (such as *fusil*, *tabac*, *femme*) in which the relationships between grapheme and morpheme are not typical and do not meet the conversion rules. If these words do not already belong to the reader's orthographic lexicon, regularization errors will occur, for example, the words will be read by generalizing the known conversion rules. For the French language, the mechanisms which need attention are the methods of using the cedilla, the cases in which /g/ and /c/ are soft or hard, or complex graphemes (Celentin 2012).

4.3 Important linguistic areas in learning of an L2

The important linguistic areas in learning of an L2 are: phonetics and phonology, morphology, vocabulary, syntax and pragmatics; for each of them it is useful to highlight possible issues or strengths that may affect students with SLD (Cappa *et al.* 2012):

- Phonetics and phonology: when we learn a foreign language, we must learn sounds that often do not exist in our native language; such sounds are not part of our phonemic inventory and can cause difficulties in both perception and production. Phonetic skills are often forgotten in language teaching, but are very important especially in the case of students with SLD. The language teacher should bring the attention of the student on similarities and differences in the sounds of the studied language to exercise the discrimination as well as production, since phonetic skills are the first that the child acquires in the native language but are not learned as spontaneously in an L2. These are issues, especially for many students with dyslexia, who usually tend to transfer into the language they are learning all the phonological difficulties they have in their native language. Learners struggle to get the phonetic models of the new language implicitly through simple exposure, so learning new sounds also requires explicit teaching. For this reason, teacher should, through targeted exercises, help the student in the discrimination and production of the sounds of the new language in order to develop a phonological awareness. It is defined as the ability to elaborate the sounds of the oral language, or the ability to analyze and manipulate the linguistic structure of words, with both appropriate verbal or non-verbal stimuli.
- Vocabulary: the most recent approaches to language teaching-learning give priority to communicative competence and give to vocabulary a central and priority role over syntactic aspects of the language. As a matter of fact, it has been stated that "l'acquisizione di un repertorio lessicale, seppur ristretto, può consentire una forma basilare di comprensione e di comunicazione, su cui potranno progressivamente inserirsi le conoscenze e competenze sintattiche" (Cappa *et al.* 2012: 10). Especially in the early stages of language learning, the teacher must suggest a particularly motivating vocabulary for students, which is as related as much as possible to their hobbies, needs, and other aspects of their daily life. In the case of students with dyslexia, the acquisition can be particularly difficult, both because of a possible deficit in the working phonological memory,[12] and because of difficulties in lexical learning due to an explicit memory deficit.[13]

[12] The phonological working memory allows to store acoustic information for a short time in order to make the correct conversion between graphemes and phonemes.

[13] Explicit memory is the one that keeps of the knowledge learned in a conscious and controlled way through study, repetition, and elaboration.

- Morphology: morphological and morphosyntactic aspects of language, as well as phonetic and phonological aspects, are learned mainly in an unconscious way in L1 and partially also in L2. Such aspects cause issues in dyslexic student in both L1 and L2. For this reason, they need to be taught in an explicit way also in L2.
- Syntax: it is important that teacher explains the differences between the syntactic constructions that may exist in different languages. Furthermore, it is essential that the consideration on syntax is preceded by a practical use of the studied structures, this must take place only after that these structures have already been assimilated as a communicative behaviour. This applies to all students but especially to the ones with dyslexia, as they may have more difficulty in recovering and using syntactic rules due to a deficiency in working memory.
- Pragmatic: as stated earlier, according to modern glottodidactics the language is understood not only as a system of rules but also as an instrument of action. Therefore language skills develop in communicative situations and are used for pragmatic purposes such as saying hello, asking for information, thanking, talking about something etc. Normally students with dyslexia have no difficulty in this area, so it is important the valorization of the pragmatic dimension in the context of language teaching-learning. Though, this may be a kind of compensation for linguistic processing difficulties. It is important that the teacher promotes more the comprehensibility of the message and the appropriateness in the communicative situation rather than assessing its correctness from the syntactic point of view. It should also be added that while there are no particular theoretical impediments to the development of socio-pragmatic skills in foreign languages, it must be considered that dyslexia can also affect the psychological level, causing frustration, depression and reluctance to communicate. Therefore, it is obvious that the development of the socio-pragmatic dimension can only take place in the presence of a serene and relaxed learning environment, in which the student is not penalized and mocked when he tries to communicate. (Daloiso 2012)

5. Problems and barrier

The dyslexic student, who is determined to learn a new language, shows a lot of curiosity and interest in learning as the foreign language is often interpreted by the learner as a second chance to demonstrate their skills. Often this initial motivation gradually shades away because of the issues the student may find on his/her path (Kormos and Kontra 2008). In losing the motivation, the student will cross more and more difficulties and, due also to the lack of the necessary supports, will tend to recreate its own barrier of protection even in the foreign language. In this path, which we can name psychological, the teacher plays a very important role in motivating the learning process; according to the emotional theory of the input of Schumann (1999, Schumann *et al.* 2004), the student evaluates the input of the teacher according to the following criteria:

- novelty: differences between the teacher's input and its own expectations;
- intrinsic pleasure: evaluation of the sense of pleasure or displeasure caused by the input, a positive evaluation helps the subject to get closer to the language to learn;
- relevance in comparison to its needs and goals: input assessment based on the obstacles or facilitation related to their social, cultural or training goals;
- viability: evaluation of the input according to its own abilities;

- psycho-social security: determine the situation according whether it can strengthen its social image or whether it harms it. (Daloiso 2012)

However, the evaluation of the input is often viewed in a negative way because the expectations are too high and are not reached in the expected time. Once the curiosity for the new language has gone away, it is much more difficult to learn it because the strategies acquired in the native tongue cannot be applied to it. It is very important for dyslexic student to reach serenity in learning new things, making the studying of the foreign language a pleasant and fun activity and not as a duty related to school success. The student must learn methods, techniques and approaches suited to his/her difficulties. Dyslexia also causes socio-affective problems, such as emotional insecurity and low self-confidence that could lead to anxiety, generally when there are challenging tests and tasks. Anxiety can be either of a character type, which is independent from the different situations, or a situational type, in other words linked to specific situations. According to some researchers, however, there is also a sort of "linguistic anxiety" which has a situational nature and it occurs especially in some cases of linguistic task such as:

- reading out in front of the whole class;
- reading out and simultaneously translating and answering questions;
- learning by heart and repeating in front of an audience;
- memorizing lists of words;
- answer questions quickly;
- improvising dialogues. (*ibidem*)

This state of anxiety can also cause a loss of communication and distraction tendencies as well as deep psychological damage; in addition, dyslexic learners tend to have a very high level of linguistic anxiety in all their activities (Kormos and Kontra 2008). As a result, it is possible to say that this state of anxiety represents the consequence of a school failure for a dyslexic student, whereas, in the case of a student without any disorder, it is the anxiety that causes school failure. Considering the obstacles related to dyslexia, the subject often thinks that the learning of a new language is extremely difficult on the cognitive level and therefore he gives up.[14] Today, thanks to the appliance of specific methods, the dyslexic student is able to reach a certain level of competence in the L2 even if some linguistic barriers could form, such as the mnestic barrier related to implicit memory.[15]

The emotional dimension pervades the entire learning environment and, for the dyslexic learner, the learning context becomes the keystone of the entire approach to teaching. If the student is emotionally involved the memorization of explicit information increases (Fabbro 2004) but, on the other hand, too strong emotional experiences can cause explicit memory systems to collapse due to the release of corticosteroid hormones that cause the blocking of the hippocampus and therefore the activation of the affective filter (Krashen in Balboni 2011). Consequently, an early intervention in the learning phases of a language allows to exploit the implicit memory and to use the language in an automatic way. For a consciously learning of words in the new language, it is necessary to activate explicit memory and ensure an appropriate

[14] It is important to clarify that speaking of cognitive level is not intended as a diminution of the abilities of the dyslexic learner facing a foreign language, but it is simply to emphasize that it is a complex process involving different dimensions simultaneously and that therefore requires to be carefully managed.

[15] The student must compensate with conscious control what he cannot automate.

environment, for example, a methodology based on *cooperative learning* (Lamberti 2010) can help the student.

The science of education of language has conceived many approaches to language teaching, the most common methods are linked to the communicative vision of the language. In Italy, Balboni's communication skills model (2008) illustrates the issues that dyslexic students might have. As far as language skills are concerned, text comprehension is a cognitive process that everyone has to carry out to understand a concept. Therefore, dyslexia does not prevent the student from reaching a good level of understanding, but can make the path more difficult by testing his/her patience and willing. Dyslexic students may also have problems in oral comprehension and production, although this is less common. In relation to levels of understanding, the dyslexic student and the student who does not have this disorder should not be put on the same level. Both of them start from a global level, pass through every other level and, finally, arrive in a deep understanding one. The only difference is that one of the two will take longer to reach the last level; it is only a matter of time and not of skills.

In accordance with the theories introduced by the functional and pragmatic linguistics, language skills develop in communicative situations and are used for precise pragmatic purposes, acquiring a socio-pragmatic character. Therefore, to reach a certain level in a language and so to be able to learn it better, the student must be able to experience it. He/she must have opportunities to use it in all actions that involve communication. In a functional way, then, it does not matter if it is grammatically correct but rather the fact of being understood by others, achieving his/her goal by being able to exchange messages. We can therefore affirm that the socio-pragmatic abilities are a communicative compensation regarding the other areas that turn out more complex.

Meta-linguistic competence has to be added to the previous skills. In the formal learning of a new language the student learns to analyze it and compares it with his/her native language searching for similarities and differences. This is a real consideration on the language, and it is one of the aspects on which the teacher must pay particular attention because the bigger problems are in the lower levels of analysis and, if ignored, they can also occur at higher levels. The student formulates and validates his/her own rules that will help him understand the language and to formulate sentences correctly. When a non-dyslexic student learns a new language, he starts by studying the morphosyntactic rules, such as word formation, tenses and verbal modes, etc. For a dyslexic student, the process takes place in a different way.

In addition, verbal language is often combined with gestures and expressions that helps to understand the language and, for this reason, we can say that a person is fully competent in a language only when he/she is able to integrate verbal language with the use of body and objects and has proper behavior. In this case, the subject must be capable to observe the different cultures and note the different behaviors used in actions.

6. Conclusions

In this text we defined the concept of dyslexia, both developmental and acquired, and we reviewed the main studies on the subject. On the biological level, we know that dyslexia is classified as a neurodevelopment disorder, with a probable genetic and hereditary basis that manifests itself clearly during the literacy process, but is already present in the subject since birth. Neuro-anatomic and neuro-physic research has tried to identify in which brain areas or neuronal circuits the disorder occurs. In the didactics some operative definitions of the disorder have been advanced, dedicated to those who are interested in language education, which

considered the different theories on dyslexia. On the other hand, on the behavioral level; there are relevant studies that have identified the signals present in different ages to highlight risk factors in order to arrive at a timely diagnosis.

In addition, we have analyzed reading and writing skills highlighting specific disorders; however, this demonstrates the need to do a deeper research into different languages and spelling systems. Research on reading learning in an L2 has focused more on the pedagogical and didactic aspects than on empirical research or theoretical models, regarding the learning and development processes of understanding in another language in learners with dyslexia.

Differences between the languages and the language learning experiences highlight that the skills involved in the written text develop differently depending on the linguistic background; therefore, the considerations of linguistic influences will lead to a better understanding of the linguistic and reading-writing performance of dyslexic learners. The identification of similarities and differences in the various linguistic contexts are necessary to understand if the predictive models referring to L1 can also be used for L2.

For this reason, it is crucial to focus more on the processes related to the consideration of the language in order to make more autonomous learners, trying to ensure the learning of L2 but also the achievement of other skills such as, for example, thinking on the strategies to be applied. Often the dyslexic student struggles with anxiety and demotivation that should not be seen as causes of dyslexia but rather are linked to school failure; an inadequate teaching can have negative consequences on the emotional, mnemonic and linguistic aspect.

References

Arfé, Barbara, Julie Rockrell and Virginia Berninger. 2014. *Writing Development in Children with Hearing Loss, Dyslexia, or Oral Language Problems*. Oxford: Oxford UP.

Associazione Italiana Dislessia. 2007. *Disturbi evolutivi specifici di apprendimento. Raccomandazioni per la pratica clinica definite con il metodo della Consensus Conference*, (Montecatini Terme, 22-23 Settembre 2006; Milano, 26 gennaio 2007), [<https://www.aiditalia.org/Media/Documents/consensus/indicazioni_consensus_DSA2007.pdf> (06/2021).]

Balboni, Paolo E. 2007. *La comunicazione interculturale*. Venezia: Marsilio.

Balboni, Paolo E. 2008. *Fare educazione linguistica. Attività didattiche per italiano L1 e L2, lingue straniere e lingue classiche*. Torino: UTET.

Balboni, Paolo E. 2011. *Le sfide di Babele. Insegnare le lingue nelle società complesse*. Torino: UTET.

Baldi, Benedetta and Leonardo M. Savoia. 2017. *Linguaggio e Comunicazione*. Bologna: Zanichelli.

Baldi, Benedetta and Leonardo M. Savoia. 2018. *Linguistica per insegnare. Mente, lingue e apprendimento*. Bologna: Zanichelli.

Camaioni, Luigia. 2001. *Psicologia dello sviluppo del linguaggio*. Bologna: Il Mulino.

Cappa, Claudia, Lucia M. Collerone, Jill Fernando and Sara Giulivi. 2014. "Apprendimento delle lingue e dislessia nella società multilingue." *Dyslang* 4: 1-41.

Cardona, Mario. 2010. *Il ruolo della memoria nell'apprendimento delle lingue. Una prospettiva glottodidattica*. Torino: UTET.

Celentin, Paola. 2012. "Didattica del francese lingua straniera ad allievi con disturbi specifici dell'apprendimento." *EL.LE* 1 (3): 605-617.

Celentin, Paola. 2016. "Interventi per l'italiano L2: problematiche e riflessioni." In *I bisogni linguistici specifici. Inquadramento teorico, intervento clinico e didattica delle lingue*, ed. by Michele Daloiso, 291-308. Trento: Erickson.

Celentin, Paola and Michele Daloiso. 2017. "La valutazione diagnostica della dislessia in contesto di L2. Problemi e riflessioni alla luce della ricerca clinico-linguistica ed interculturale." *EL.LE* 6 (3): 323-247.

Chiappetta Cajola, Lucia and Marianna Traversetti. 2017. *Metodi di Studio e DSA. Strategie didattiche inclusive*. Roma: Carocci Faber.

Coltheart, Max. 1981. "Disorder of Reading and Their Implication for Models of Normal Reading." *Visible Language* 15 (3): 245-286.
Cornoldi, Cesare (a cura di). 2007. *Difficoltà e disturbi dell'apprendimento*. Bologna: Il Mulino.
Costenaro, Verusca, Michele Daloiso and Luciana Favaro. 2014. "Teaching English to Young Learners with Dyslexia: Developing Phonemic Awareness Through The Sound Pathways." *EL.LE* 3 (2): 209-230.
Cummins, Jim. 1979. "Cognitive/Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters." *Working Papers on Bilingualism* 19: 121-129.
Daloiso, Michele. 2012. *Lingue straniere e dislessia evolutiva. Teoria e metodologia per una didattica accessibile*. Torino: UTET.
Defior, Sylvia, Francisco, Martos and Luz, Cary. 2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. Applied Psycholinguistics, 23: 135-148.
Dehaene, Stanislas. 2007. *Les neurones de la lecture*. Paris: Odile Jacob.
Dehaene, Stanislas. 2009. *Reading in the Brain. The New Science of How We Read*. New York, NY: Viking Penguin.
Duncan, Lynn G., Pascale Colé, Philip H.K. Seymour, and Annie Magnan. 2006. "Differing Sequences of Metaphonological Development in French and English." *Journal of Child Language* 33 (2): 369-399. doi: 10.1017/S030500090600732X.
Fabbro, Franco. 2004. *Neuropedagogia delle lingue. Come insegnare le lingue ai bambini*. Roma: Astrolabio.
Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge: CUP.
Frost, Ram and Leonard Katz (eds). 1992. *Orthography, Phonology, Morphology and Meaning*. Amsterdam: Elsevier.
Galaburda, Albert M. (ed.). 1989. *From Reading to Neurons*. Cambridge: MIT Press.
Galaburda, Albert M. 2005. "Dyslexia – A Molecular Disorder of Neuronal Migration. The 2004 Norman Geschwind Memorial Lecture." *Annals of Dyslexia* 55 (2): 151-165. doi: 10.1007/s11881-005-0009-4.
Geschwind, Norman and Albert M. Galaburda. 1985. "Cerebral Lateralization: Biological Mechanisms, Association and Pathology: A Hypothesis and a Program for Research." *Archives of Neurology* 42 (5): 428-459.
Grech, Louisa, Amir Sadeghi and John Everatt. 2016. "Valutare le difficoltà di comprensione del testo scritto in apprendenti plurilingue in contesto educativo." In *I bisogni linguistici specifici. Inquadramento teorico, intervento clinico e didattica delle lingue*, ed. by Michele Daloiso, 167-182. Trento: Erickson.
Juan Eugenio Jiménez González. 2012. Dislexia en español: prevalencia e indicadores cognitivos, culturales, familiares y biológicos. Ediciones Pirámide: Madrid.
Kormos, Judit and Edit H. Kontra (eds). 2008. *Language Learners with Special Needs: An International Perspective*. Bristol-Buffalo-Toronto: Multilingual Matters.
Krashen, Stephen D. and Tracy D. Terrel. 1983. *The Natural Approach. Language Acquisition in the Classroom*. Hayward, CA: Alemany Press.
Lamberti, Stefania. 2010. *Apprendimento cooperativo e educazione interculturale*. Trento: Erickson.
Matute, Esmeralda and Soledad Guajardo. 2012. *Dislexia. Definición e intervención en hispanohablantes*. México: Editorial El Manual Moderno.
Melero Rodríguez, Carlos A. 2015. "Dislessia, glottodidattica speciale e 'scienze del cervello' ". In *Glottodidattica e neuroscienze: verso modelli traslazionali*, ed. by Giovanni Buccino and Marco Mezzadri, 187-198. Firenze: Franco Cesati.
Ministero dell'Istruzione, dell'Università e della Ricerca. 2012. *Direttiva Ministeriale 69 1 agosto 2012. Linee guida per i percorsi degli Istituti Tecnici*. <www.gazzettaufficiale.it/eli/id/2012/10/29/12A11381/sg> (06/2021).
Ministero dell'Istruzione, dell'Università e della Ricerca. 2012. *Direttiva Ministeriale 70 1 agosto 2012. Linee guida per i percorsi degli Istituti Professionali*. <www.gazzettaufficiale.it/eli/id/2012/10/29/12A11382/sg> (06/2021).
Ministero dell'Istruzione, dell'Università e della Ricerca. 2011. *Linee guida per il diritto allo studio degli alunni e degli studenti con disturbi specifici dell'apprendimento*. <https://www.miur.gov.it/disturbi-specifici-dell-apprendimento-dsa-> (06/2021).
Nelson, Katherine. 1986. *Event Knowledge: Structure and function in development*. Hillsdale, NJ: Laurence Erlbaum Associates, Inc.

Nicolson, Roderick I. and Angela J. Fawcett. 2008. *Dyslexia, Learning and the Brain*. Cambridge, MA-London: The MIT Press.

Paulesu, Eraldo, Elisa Samburago, T. Torti, L. Danelli, F. Ferri, G. Scialfa, M. Sberna, G.M. Ruggiero, G. Bottini, S. Sassaroli. 2010. "Neural Correlates of Worry in Generalized Anxiety Disorder and in Normal Controls: A Functional MRI Study." *Psychological Medicine* 40 (1): 117-124. doi: 10.1017/S0033291709005649.

Schumann, John H. 1999. *The Neurobiology of Affect in Language*. Oxford: Blackwell.

Schumann, John H., Sheila E. Crowell, Nancy E. Jones, Namhee Lee, and Sara Ann Schuchert. 2004. *The Neurobiology of Learning. Perspectives from Second Language Acquisition*. London: Routledge.

Seymour, Philip H.K., Mikko Aro and Jane M. Erskine. 2003. "Foundation Literacy Acquisition in European Orthographies." *British Journal of Psicholo*gy 94 (2): 143-174. doi: 10.1348/000712603321661859.

Sabbadini, Giorgio. 1995. *Manuale di neuropsicologia dell'età evolutiva*. Bologna: Zanichelli.

Smythe, Ian. 2004. *Provision and Use of Information Technology with Dyslexic Student in University in Europe*. <https://eprints.soton.ac.uk/264151/1/The_Book.pdf> (06/2021).

Snowling Margaret. 1987. *Dyslexia. A Cognitive Developmental Perspective*. Oxford: Blackwell.

Stein, John F. 2001. "The Magnocellular Theory of Developmental Dyslexia." *Dyslexia* 7 (1): 13-36. doi: 10.1002/dys.186.

Stein, John F. 2003. "Visual Motion Sensitivity and Reading." *Neuropsychologia* 41 (13): 1785-1793. doi: 10.1016/S0028-3932(03)00179-9.

Stella, Giacomo. 2004. *La dislessia*. Bologna: Il Mulino.

Talcott, Joel B., Peter C. Hansen, Charles Willis-Owen, Ian W. McKinnell, Alex J. Richardson, and John F. Stein. 1998. "Visual Magnocellular Impairment in Adult Development Dyslexics." *Neuro-Ophthalmology* 20 (4): 187-201. doi: 10.1076/noph.20.4.187.3931.

Talcott, Joel B., Peter C. Hansen, Elikem L. Assoku, and John F. Stein: 2000. "Visual Motion Sensitivity in Dyslexia: Evidence for Temporal and Energy Integration Deficits." *Neuropsychologia* 38 (7): 935-943. doi: 10.1016/S0028-3932(00)00020-8.

Triolo, Riccardo. 2016. "Interventi specifici per le lingue classiche." In *I bisogni linguistici specifici. Inquadramento teorico, intervento clinico e didattica delle lingue*, ed. by Michele Daloiso, 273-290. Trento: Erickson.

Zanobini, Mirella and Maria C. Usai. 2008. *Psicologia dell'handicap e della riabilitazione: i soggetti, le relazioni, i contesti in prospettiva evolutiva*. Milano: Franco Angeli.

Ziegler, Johannes and Usha Goswami. 2005. "Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory." *Psychological Bulletin* 131 (1): 3-29. doi: 10.1037/0033-2909.131.1.3.

Wolf, Maryanne and Patricia G. Bowers. 1999. "The Double-Deficit Hypothesis for the Developmental Dyslexia." *Journal of Educational Psychology* 91 (3): 415-438. doi: 10.1037/0022-0663.91.3.415.

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: Sergio Conti (2021) Negotiation in Chinese as a foreign language: Italian learners' signal strategies in a tandem-learning context. *Qulso* 7: pp. 289-313. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12012

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Negotiation in Chinese as a foreign language: Italian learners' signal strategies in a tandem-learning context

Sergio Conti
Università degli Studi Roma Tre (<sergio.conti@uniroma3.it>)

Abstract:

Negotiation of meaning, the process through which the speakers go to clearly comprehend one another, is believed to facilitate acquisition as it provides learners with situated comprehensible input as well as opportunities to produce pushed output. Nevertheless, negotiation is still an under-researched area in Chinese as a second or foreign language. This study analysed the strategies used by learners to signal non-understanding and the linguistic resources they exploit to achieve this goal. The participants were 13 Italian learners who took part in a three-month face-to-face tandem language learning program with 6 native speakers of Chinese. The interactions were recorded and transcribed, and the produced signals were classified according to strategy, form, and language type. The sequential position of strategies in the negotiation routines was also annotated. Empirical evidence shows a strong preference for repetitions and a frequent code-switching to Italian. A clear preference in the sequential distribution of clarification requests and confirmation checks also emerged, the former being mainly used to initiate negotiations routines, and the latter mostly occurring in routine reiterations.

Keywords: *Chinese a Foreign Language, Clarification Requests, Comprehension Checks, Negotiation of Meaning, Tandem Language Learning*

1. Introduction

According to Long's (1985, 1996) Interaction Hypothesis, negotiation of meaning (NoM) in conversational interaction provides the optimal conditions for input to become comprehensible. Interactionally-motivated modifications of the input facilitate language acquisition as learners' attentional resources are oriented to the mismatches between what they know and what emerges from interaction, as well as areas of the L2 about which they have little or no information. In addition, NoM also promotes the production of pushed output, in that, after

receiving negative feedback, learners are encouraged to reformulate their message by trying out new structures (Swain 1985, 2000). In sum, NoM has the twofold potential of assisting comprehension and providing opportunities to test L2 hypotheses.

Due to these reasons, NoM has received a fair amount of attention in Second Language Acquisition. Researchers have explored various aspects of NoM, including its effects on comprehension and production (e.g., Gass and Varonis 1994) and the interplay of factors such as gender (e.g., Pica *et al.* 1991), age (e.g., Oliver 2002) task type (e.g., Pica and Doughty 1985; Nakahama, Tyler, and van Lier 2001), and learning context (e.g., Fernández-García and Martínez-Arbelaiz 2002; Gass, Mackey, and Ross-Feldman 2005; O'Rourke 2005). Overall, research has convincingly demonstrated a relationship between NoM and L2 learning (see Pica 1994). Yet, NoM in Chinese as a second (CSL) or foreign language (CFL) is still an under-researched area.

In order to fill this gap, this study will investigate how Italian CFL learners engage in negotiation during the interactions with native peers in a face-to-face tandem-learning context. In particular, the analysis will focus on the pragmatic strategies used by learners to signal comprehension problems and the linguistic resources they exploit to achieve this goal.

2. Literature review

2.1 Negotiation of meaning: A model

NoM refers to the work that speakers do to avoid and repair impasses in their conversational discourse (Long 1983). It consists in interactional modifications resulting from communication breakdowns, aimed to achieve mutual understanding (Nakahama, Tyler, and van Lier 2001: 378). In this sense, NoM is a kind of linguistic problem-solving for learners, in that, "[t]o repair the breakdown, the interlocutors must somehow negotiate the disparity between their L2 abilities and the target language" (Doughty 2000: 48).

Note that NoM is not limited to the interactions involving second language learners. Indeed, this process is much more general in the normal use of language, as it reflects an essential property of natural languages and their use, i.e. the vagueness of meaning and in referring to the world, as well as the linguistic differences between speakers. Bazzanella and Damiano (1997), for instance, distinguished five levels of misunderstanding in a corpus of everyday conversations between native speakers of Italian. These levels correspond to the five levels of linguistic analysis – phonetics, syntax, lexis, semantics, and pragmatics. In a subsequent study (Bazzanella and Damiano 1998), the authors further analysed the phases of the negotiations and the collocations of the repairs, and concluded that comprehension should be interpreted as a *continuum* rather than a polar process (comprehension/non-comprehension). Another study is that by Comeau, Genesee, and Mendelson (2010), who compared the repair skills of bilingual and monolingual children. The authors found no differences between the two samples, which led to the conclusion that "the acquisition of this communication skill is fundamentally unperturbed by the simultaneous acquisition of two languages in the first years of life" (*ibidem*, 371).

Going back to second language interactions, the pivotal work by Varonis and Gass (1985) proposed a model for NoM that described negotiation routines as consisting of two parts and four functional primes. The first part of the routine is the trigger (T), which entails one prime with the same name. It consists in the utterance or portion of the utterance that causes non-understanding and can initiate from any aspect of discourse. The trigger is only recognised in retrospect, that is, if it has been reacted to by the hearer.

The second part of the sequence, the resolution, entails three primes, an indicator or signal, a response, and a reaction to the response. The signal (S) refers to the hearer's observable (verbal or non-verbal) reaction to the trigger, which prompts the speaker to repair the problematic utterance. The speaker's response (R) to the signal can then be followed by an optional reaction (RR), which ties up the routine before returning to the main flow of conversation.

This analytical model was applied in numerous studies, including those on CFL/CSL (e.g., Wang 2006; Wang and Li 2015; see § 2.2). Nakahama, Tyler, and van Lier (2001), for instance, investigated how meaning is negotiated in two different types of task, unstructured conversations and a two-way information-gap task, and analysed the NoM routines in terms of type of trigger, length of turns, complexity of utterances, and discourse strategies. According to the results, unstructured conversations "provided the NNS interlocutors with more opportunities to hear more complex input from the NS interlocutors and with more opportunities to produce syntactically complex output" (*ibidem*, 391). NNS utterances in unstructured conversations were also longer and more complex, and the discourse strategies adopted were extremely various, including hedges, reformulations, and demonstrations of understanding. Lastly, the information-gap task was characterised by chained NoM sequences, with NoM only occurring locally; in conversations, on the other hand, NSs and NNSs were found to negotiate meaning in order to achieve coherence in the entire interaction.

More recently, Van der Zwaard and Bannink (2019) analysed NoM following a microanalytical approach. They examined the online interactions between six NS-NNS dyads and found that, in tasks with multiple triggers of potential misunderstanding, NS tended to use responses to signals that were both task- and face-appropriate, whereas the NNS gradually moved from task- to face-appropriate responses.

Lastly, the study by Lee, Hampel, and Kukulska-Hulme (2019) applied the NoM model to examine the role of gestures during videoconferencing. According to their analysis, iconic and deictic gestures were used to establish mutual understanding and negotiate vocabulary, operating in conjunction with speech to trigger NoM and serving as a resource with which to scaffold peers and indicate appeals for assistance. However, gestures were also a possible source for confusion, contributing to further incidents of non-understanding.

What is more relevant for the present study is the second prime of Varonis and Gass's sequence (1985), i.e., the indicator or signal. A first formal taxonomy was proposed by Varonis and Gass (*ibidem*), who distinguished between explicit indications of non-understanding (mostly *wh-* questions and statements such as 'I don't understand'), echoes (repetitions), nonverbal responses, and inappropriate responses.

Other studies identified the pragmatic strategies used to indicate non-understanding. These strategies have to do with the signaler's utterance planning and intentionality (Pica 1994: 497), and the most common types are clarification requests (CRs) and confirmation checks (CCs).

CRs are used to elicit clarification of the preceding utterance (Long 1980). They occur when one speaker does not understand one or more words being articulated by the interlocutor and thus needs re-explaining, reuttering, or rephrasing (Masrizal 2014). They typically consist in *yes-no* or *wh-* questions (e.g., 'What do you mean?'), statements inviting the interlocutor to repeat, or explicit statements of misunderstanding. By contrast, CCs are defined as "[e]licitations immediately following the previous speaker's utterance to confirm that the utterance has been understood or heard correctly" (Chen 2016: 6; see also Long 1980). Typical manifestations of CCs are summaries or reformulations (e.g., 'Do you mean he has something on his head?') and full or partial repetitions of the preceding utterance (Chen 2016: 7).

These labels have been applied extensively in the literature (e.g., Masrizal 2014). Lee (2001), for instance, investigated the strategies employed by 40 learners of Spanish to facilitate

negotiation during online discussions. She analyzed 289 signals, observing a predominance of CRs (19%), requests (20%), and self-corrections (16%). Gass, Mackey, and Ross-Feldman's (2005) study involved 74 learners of Spanish. The authors compared interactions in classrooms and laboratories, finding that both conditions triggered a similar amount of CRs and CCs. A significant difference for both strategies was observed among different tasks. Chen (2016) examined the strategies used by 9 adult learners of English during task-based interactions in Second Life and found a prevalence of CCs (41%) over CRs (33%).

A very fine-grained description of signals is that conducted by O'Rourke (2005) in a study on NoM in computer-mediated tandem learning. The participants in the study were Irish and German learners of German and English, respectively. The author described both the pragmatic strategies adopted by the participants and the form in which these strategies manifested themselves. He identified a total of 95 signals and found that many consisted in hybrid forms, carrying the illocutionary forms of both CRs and CCs. An example is the signal *Wie du bist positiv. HIV-positiv?* 'How do you mean you're positive. HIV-positive?' (*ibidem*, 445), a combination of a *wh-* question (CR) and a summary/reformulation (CC).

2.2 Negotiation in Chinese

According to Su and Hu (2017: 75-76), Chinese researchers have started to devote attention to negotiation only recently, and the limited number of contributions often failed to bring out CFL/CSL specificities (e.g., Qian 2010).

The study by Wang and Li (2015) constitutes an exception. The authors explored the turn-taking structure of negotiation routines in naturalistic interactions between a native (NS) and a non-native speaker (NNS) of Chinese, finding a series of statistically significant differences between the two. The NS, for instance, produced 76 signals, while the NNS only produced 25. The main strategy to indicate non-understanding consisted of CCs, with the NS showing a strong preference for reformulations of the trigger embedded in conversational chunks (e.g., *nǐ shì xiǎng shuō* 你是想说…… 'you mean…'). By contrast, the NNS tended to produce simple repetitions in a rising intonation (e.g., *cóng* 丛？ 'clump?'). As for CRs, the NS often resorted to *wh-* questions, while the NNS mostly used the interrogative marker *á* 啊？ 'uh?' with little variation. For both speakers, the majority of signals consisted of simple forms, with hybrid forms only playing a marginal role.

The authors also analyzed the form of the responses and reactions. The responses provided by the NNS mostly consisted of expansions, acknowledgements (e.g., *duì* 对 'right'), or negations, while the NS's responses mainly consisted of repetitions. Lastly, both the NS and the NNS's reactions were produced in the form of response markers (e.g., *hǎo* 好 'good') or topic expansions.

A second example is the study by Zhao (2015), who analyzed 30 intermediate-level learners' negotiation routines in two different tasks, an information gap task and an opinion gap task. The author found that the two tasks led to the same amount of negotiation routines but differed in the type of strategy used to signal non-understanding. In the information gap tasks, the participants produced a higher percentage of CCs, while in the opinion gap task they showed a clear preference for CRs. The author also observed that requests for help were the only strategy used to negotiate form in opinion gap tasks. According to Zhao, these results partially contradict previous studies, but this might depend on factors such as the design and requirements of the tasks and the proficiency level of the participants.

Another group of studies focused on the interactions in a virtual environment. Wang (2006), for instance, documented NoM in computer-mediated interactions between 8 CFL

learners and their teacher. The author found instances of all four kinds of indicators described by Varonis and Gass (1985), with explicit statements of non-understanding being the majority. Cappellini and Zhang (2013) adopted a multimodal approach to describe NoM routines in teletandem interactions between a French learner of Mandarin and a Chinese peer. According to their analysis, the signals produced by the French learner mainly consisted of CCs (39.6%) and requests for help (32.1%), whereas triggers were mostly involved vocabulary. Overall, the authors observed that teletandem favored the diversification of NoM modalities (e.g., written and spoken), and that the number of NoM routines was higher than other studies set in different interactional environments.

Given the paucity of contributions on NoM in CFL, this study aims to shed more light on this issue by describing negotiation routines in NS-NNS interactions, with particular focus on the pragmatic strategies used by the NNS to signal comprehension problems.

The interactions took place during a face-to-face tandem language-learning program. Tandem is a method of language learning in which speakers of different first languages (L1) communicate regularly with one another, each wishing to learn the interlocutor's L1. It is based on the two principles of autonomy and reciprocity, as participants are at the same time responsible for their own as well as their partner's learning (Brammerts 2003). Thus, tandem promotes authentic, culturally grounded interaction, while also combining explicit form-focused learning and meaningful communication. Crucially, discourse in a tandem-learning context is highly naturalistic, offering fertile ground for NoM. In addition, tandem partners are free to choose between conversational and pedagogical topics according to their needs; therefore, meta-linguistic information is highly targeted. This provides optimal conditions to exploit the benefits of comprehensible input and pushed output as theorised by the Interaction Hypothesis.

Based on these premises, the research questions (RQs) of this study are the following:

1. What are the pragmatic strategies adopted by CFL learners to signal comprehension problems and start negotiation routines in a tandem-learning context?

2. What is the form of the signal, i.e., what linguistic resources do learners exploit to indicate non-understanding?

3. What is the role of the participants' L1?

3. Method

3.1 Participants

The participants were 13 Italian second-year CFL learners who voluntarily joined the Italian-Chinese tandem program at the Roma Tre University Linguistic Centre (CLA). The native partners were 6 Chinese learners of Italian, who were part of the Marco Polo/Turandot program at CLA.[1] Although not measured, it is safe to assume that the competence level of the two groups of participants in their respective L2 was roughly equivalent. In fact, Marco Polo/

[1] The Marco Polo/Turandot program aims at providing Chinese students with proficiency in Italian language and culture and preparing them for enrolment at Italian Universities, Academics of Fine Arts, and Conservatories.

Turandot students are required to obtain a certification of Italian language proficiency corresponding to level B1 of the Common European Framework for Languages (CEFR). Similarly, the target level for second year at Roma Tre University corresponds to level 3 of the Chinese Proficiency Test (*Hanyu Shuiping Kaoshi* 汉语水平考试), which corresponds to level B1 in the CEFR (but see Lu 2017). All the Italian participants were females, whereas the Chinese participants were one male and five females. All the participants signed an informed consent for participation in the study.

3.2 Data collection

The data were collected throughout a three-month timespan, during which the participants met at CLA once a week. Due to the numerical disparity of the two groups, each Chinese speaker was grouped with one or two Italian partners. The duration of each encounter was about 1 hour. A moderator signaled the switch from Italian to Chinese and vice versa at regular intervals of about 5-10 minutes.

The data were collected during six encounters, equally distributed throughout the three months of observation. The interactions in Chinese were recorded using a mobile device connected to a unidirectional noise-cancelling microphone with -46 ± 2 dB sensitivity and 50-16,000 Hz frequency band. The wire connecting the microphone to the device was 3 m long, and this minimized the possibility of the researchers' presence inhibiting the participants. Due to technical constraints, it was only possible to record one group at the time. The total length of the recordings was approximately 2h50', with a mean length of 24' per encounter and 6'22" per recording. The audio data were transcribed by trained transcribers and then revised by the researcher. The transcriptions follow the Jeffersonian conventions for Conversation Analysis (see Jefferson 2004; Hepburn and Bolden 2013), adapted to Chinese following Lu, Lee, and Tao (2014) and Lee, Tao, and Lu's (2017) suggestions.

3.3 Analysis

Negotiation routines were identified in the transcripts and coded according to language (Italian or Chinese), strategy type, routine type, and the sequential position of strategies.

Firstly, signals were coded as "Italian" or "Chinese" according to the language of the utterance. Some signal forms were necessarily in Italian (e.g., translations) or Chinese (e.g., repetitions), while other signals occurred in both languages (e.g., explicit statements of non-understanding). Note that the annotation is based on the predominant language of the utterance, even if elements in the other language are also present. An example is the CC 'simile a *yìdiǎnr* 一点儿' (similar to *yidianr*). In this case, the NS proposed a reformulation (*yìdiǎnr*一点儿 'a little') to the trigger *bù zěnme* 不怎么 'not that much', but the utterance in which it was embedded was clearly Italian. Lastly, signals that did not display a clear preference for any of the two languages, including silent signals, were all labelled as "unclear".

Secondly, the classification of strategy types followed a data-driven approach. In other terms, it was not determined a priori based on formal features associated to pre-existing categories. Instead, it was contextually determined, allowing for features of interaction to describe themselves. To do so, the Conversation Analytic technique of the next turn-proof procedure was used. According to this technique, the interlocutor's response to an utterance displays their understanding of the prior turns' talk and can be taken as evidence for what action was accomplished in conversation (Sacks, Schegloff, and Jefferson 1974). Following this procedure, the reaction to the signal was used to determine the category in which the strategy was classified.

Repetitions of the trigger constitute a particularly explicative case. Previous studies (e.g., Chen 2016: 7) considered repetitions as one of the typical realizations of CCs; however, this categorization disregards the way in which interlocutors co-construct meaning during interaction. To clarify this point, compare the examples reported in Excerpts 1-2.[2]

(1) Day2_#02

37	CH2:	e	什么类型	(1.0)	>类型<		
		e	*shénme*	*lèixíng*	*(1.0)*	*>lèixíng<*	
		PRT	what	kind		kind	
		‘Uh what kind (1.0) >kind<’					
38 →	IT5:	l- lei sing?					
39	CH2:	m tippo: (.) tippo in cinese:^ è 类 [型.					
		m	*tippo: (.)*	*tippo*	*in*	*cinese:^ è*	*lèi [xíng.*
		PRT	kind	kind	in	Chinese is	kind
		‘Mh %ki:nd% (.) %kind in Chine:se^ is% *lei [xing.*’					

(2) Day2_#05

71	CH4:	… 所以我必须每天都要做饭.							
		… suǒyǐ	*wǒ*	*bìxū*	*měi*	*tiān*	*dōu*	*yào*	*zuòfàn.*
		so	1sg	must	every	day	all	must	cook
		‘… so I must cook every day.’							
72 →	IT10:	做饭.							
		zuòfàn.							
		cook							
		‘*Zuofan*’							
73	CH4:	就是.							
		jiù	*shì.*						
		just	be						
		‘That’s right.’							

In Excerpt 1, IT5 attempted the inaccurate question *shénme yàngzi* 什么样子 ‘what model?’ to ask CH2 what kind of music they liked. CH2 provided corrective feedback by saying that the appropriate term was *lèixíng* 类型 ‘kind’. Clearly, IT5 did not know this word, as suggested by her attempts to mimic its pronunciation (line 38). The evident hesitations and rising intonation suggest that the underlying illocutionary force of the utterance was asking for clarification. This is confirmed by CH2’s prompt explanation in line 39. Following the next-turn proof procedure, CH2’s response can be used as evidence for interpreting IT5’s signal as a CR.

[2] In the excerpts, ‘#’ indicates the progressive number of the recording in that specific day. Italian participants are indicated through the notation ‘IT’ followed by an identification number from 1 to 13. Similarly, Chinese participants are indicated through the notation ‘CH’ followed by an identification number going from 1 to 6. Interlinear glosses follow the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>(06/2021), adapted to Chinese based on Li and Thompson’s (1981) notation. The abbreviations ACT and PRT are not included in the cited references: the former stands for ‘actualising *ci*’ (‘*ci* attualizzante’), a particular use of the Italian locative pronoun ‘ci/ce’ (see Sabatini 1985: 160); the latter stands for ‘particles’ and refers to various types of non-lexical vocalisations, including exclamations, interjections, backchannels, hesitations, etc. (e.g., see Wu 2016). These particles were all transcribed in Latin letters, as most of them cannot be written in conventionalised Chinese characters (Lee, Tao, and Lu 2017: 792).

In Excerpt 2, CH4 told IT10 that she had to cook for her boyfriend every day, using the word *zuòfàn* 做饭 'to cook'. This was probably unknown to IT10; nonetheless, it was easily inferable. In fact, *zuò* 做 'to make' is a very common verb, and *fàn* 饭 'cooked rice' is the general term to express 'meal', as in the compound *chīfàn* 吃饭 [eat-rice] 'to eat'. Contrarily to Excerpt 1, the repetition in line 72 is uttered in a falling intonation and without hesitations. This suggests that IT10 repeated the word to confirm her inference, and this is consistent with CH4's acknowledgement in line 73 ('that's right'). Thus, the repetition here can be interpreted as a CC.

In summary, the two excerpts show that two apparently identical strategies are in fact different in nature, as suggested by the NSs' reactions. Additional interpretive evidence is provided by features of the speaker's vocal conduct such as intonation, hesitations, phonetic accuracy, etc.

Once pragmatic strategies were identified, sub-categories were created according to the form of the signal, i.e., the linguistic resources the participants exploited to perform such strategies.

A further distinction was made between simple and embedded routines. Simple routines are those following the typical T-S-R-RR sequence, whereas embedded or 'nested' routines occur when the incomprehension persists after the R (Varonis and Gass 1985: 78; Doughty 2000: 48). In the latter case, the NS's response is followed by a new signal instead of the reaction. This reiterates the routine until comprehension is reached or negotiation is abandoned.

Lastly, NNS's strategies to signal non-understanding were classified at the sequential level, distinguishing between those occurring in routine initiation and those occurring in subsequent reiterations after the R.

3.4 Results

A total number of 78 NS-triggered negotiation routines were identified in the transcripts. Of these, 46 (58.97%) were simple routines and 32 (41.03%) were embedded or "nested" routines.

As noted by Varonis and Gass (1985: 78-81), embedded routines can display multiple layers and entail more than one trigger. In Excerpt 3, CH6's attempts to clarify the word *Wànshénmiào* 万神殿 'Pantheon', a famous monument in Rome, generated new understanding problems (lines 11, 16, and 24). The original problem was only resolved in turn 33, after CH6 tried to translate the word into Italian.

(3) Day3_#06

7	CH6	… 在: (.) 万神殿附近.
		… *zà:i (.)* *Wànshénmiào* *fùjìn.*
		be.at Pantheon nearby
		'… It's: (.) near the Pantheon.'
8 →	IT11	uan shi?
9	CH6	万神殿 °附近° (1.0) e:: 一个 (0.5) <圆顶> (.) (这边) 有一个洞子.
		Wànshénmiào °fùjìn° (1.0) *e::* *yī-gè (0.5)* *<yuándǐng> (.)*
		Pantheon nearby VCL one-CLF round-top
		(zhèbiān) *yǒu* *yī-gè* *dòngzi.*
		here have one-CLF hole
		'°Near° the Pantheon (1.0) uhm:: one ((with a)) (0.5) <round top> (.) (here) has a hole.'
10	IT11	=AH era grande (0.2) 大的:: (0.7) 很大.
		=AH *era* *grande (0.2)* *dà-de:: (0.7)* *hěn* *dà.*
		PRT was.2sg big big-DET very big
		'=%AH it was big% (0.2) bi::g (0.7) very big.'

11 CH6 (not clear) >很大一个< (.) 圆的: (.) 房顶.
(not clear) >hěn dà yí-gè< (.) yuán-de: (.) fángdǐng.
very big one-CLF round-DET house-top
'(not clear) >a very big one< (.) ((with a)) round: (.) roof.'
((omitted))
16 → IT12 fanding?

17 IT11 la coppetta?
the ice-cream.cup
'%The ice-cream cup?%'
18 CH6 =等一下>在万神殿附近是一个< (.) 经典 monumendo (1.0) eh: vicino a un mo-
=děng yíxià >zài Wànshénmiào fùjìn shì yī-gè< (.)
wait one-moment be.at Pantheon nearby be one-CLF
jīngdiǎn monumendo (1.0) e:h vicino a un mo-
classic monument PRT nearby at a mo(nument)
'=Wait >near the Pantheon it's a< (.) classic %monument (1.0) e:h near a mo-%'
19 IT11 [AH VICINO A UN MONUMENTO hhh.h
PRT near at a monument
'[%AH NEAR A MONUMENT hhh.h%'
((omitted))
21 CH6 然后这个 (.) e monumento 是一个 (.) 圆顶的,
ránhòu zhè-ge (.) e monumento shì yī-ge (.)
then this-CFL PRT monument be one-CFL
yuándǐng-de,
round-top-NOM
'Then this (.) uh %monument% has a round top,'
((omitted))
24 → IT12 [yuanding che cos'è?
yuanding what thing-be
'[What's *yuanding*?'
((omitted))
29 CH6 Ah 这个°pa-° (1.0) °pa-° (0.7) <passone?> pansone?
ah zhè-ge °pa-° (1.0) °pa-° (0.7) <passone?> pansone?
PRT this-CLF
'Ah this °pa-° (1.0) °°pa-°° (0.7) <passone?>pansone?'
((omitted))
33 IT12 Pia- ah ^Pantheo:n:?

Overall, 154 pragmatic strategies were used by the non-native participants to signal communication breakdown. Table 1 summarizes the overall frequency and percentage of each type of pragmatic strategies that occurred in the data. In addition, Table 1 also illustrates the sequential organization of strategy use, by reporting the percentage of occurrence of each strategy type in routine-initiating position and in subsequent reiterations.[3]

[3] In Table 1, the number of routine initiations (83) is superior to the total number of negotiation routines (78) because in five cases, the routine was initiated by both Italian participants in the same tandem group.

			Initiations		Reiterations	
Strategy type	**N**	**%**	**N**	**%**	**N**	**%**
Clarification requests	66	42.86%	42	50.60%	24	33.80%
Confirmation checks	61	39.61%	22	26.51%	39	54.93%
Implicit signals	14	9.09%	13	15.66%	1	1.41%
Other strategies	13	8.44%	6	7.22%	7	9.86%
Total	154	100%	83	100%	71	100%

Table 1. Frequency and percentage of strategy use

Empirical observation shows that the great majority of the strategies used by the NNS participants to initiate negotiation routines consisted of CRs and CCs. Due to their overall predominance, these two categories will be discussed more extensively in Section 3.5.

A considerably smaller percentage (9.09%) consisted of implicit signals (IS), which include silence and a variety of pragmatic markers. Examples of IS are reported in Excerpts 4-6. Excerpt 4 is an example of silent signal. CH2 asked IT3 if she had a boyfriend before, but IT3 remained silent. Note that the response to the signal was provided by the other NNS (line 127). In Excerpt 5, the NNS's hesitation served as an IS that initiated the routine. Again, the response was provided by the other non-native interlocutor. In Excerpt 6, the NNS used the backchannel 'mhmh' (line 2), which prompted the NS to produce the subsequent comprehension check (line 25). Backchannels are signals that indicate that the interlocutor can continue talking and are often used to elicit more conversation or elucidation (De Bartolo 2014: 458). In Excerpt 6, it appears that the NNS's backchannel was interpreted by the NS as an incomprehension signal. Seemingly, this interpretation was correct, as suggested by the NNS's reactions to the NS's responses (lines 26, 28).

(4) Day2_#02

125 CH2: … 以前^ (0.5) 以前有吗？
… yǐqián^ (0.5) yǐqián yǒu ma?
before before have Q
'… Before^ (0.5) Did ((you)) have ((one)) before?'

126 → IT3: (0.7)

127 IT5: Prima ce l' avevi?
Before ACT 3sg.OBJ had-2sg.
'Did you have one before?'

((omitted))

129 IT3: 一定.
Yídìng.
Surely
'Sure.'

(5) Day3_#01

16 CH1: 所以你是在 (0.2) 九七年 (.) 出生的.
suǒyǐ nǐ shì zài (0.2) jiǔqī nián (.) chūshēng-de.
so 2sg be in nine-seven year be.born-MOD

'So you were born in ((nineteen)) ninety seven?'

17 → IT3: E::

18 (1.0)

19 IT6: °Novantasette.°
Ninety-seven
'%°Ninety-seven.°%'
20 IT3: =Ah sì.
PRT yes
'=%Oh yeah.%'

(6) Day1_#02
23 CH4: 我妹妹已经工作了 (.) 她是一名 (.) 服装设计师.
wǒ mèimei yǐjīng gōngzuò le (.)
1sg little.sister already work CRS
tā shì yī míng (.) fúzhuāng shèjìshī.
2sg.F be a CL dress designer
'My little sister already works (.) She's a (.) dress designer.'
24 → IT3: Mhmh^
25 CH4: 你知道什么是服装设计师吗？ E:: (.) visto (.) °visto°?
nǐ zhīdào shénme shì fúzhuāng shèjìshī ma?
2sg know what be dress designer Q
e:: (.) visto (.) °visto°?
PRT dress dress
'Do you know what a *fuzhuang shejishi* is? U::h (.) visto (.) °visto°?'
26 IT3: aha^
27 CH4: de- (.) desi:gn.
28 IT3: ok.

Lastly, a heterogeneous variety of strategies that did not fall into any of the previous categories was grouped under the label "other strategies" (OS). In Excerpt 7, for instance, the NNS used a meta-cognitive strategy, consisting in the speakers' verbalization of her thinking process (Chen 2016: 11). Due to the exiguous number of occurrences, this category will not be discussed further.

(7) Day2_#04
3 CH1: [HH 没准备好.
[HH *méi zhǔnbèi-hǎo.*
PRT NEG prepare-well
'Ahah (you're) not ready.'
4 IT7: °spe che ha chiesto che (sta a di)°?
wait what has asked what (is PROG say)
'°Wait what has he asked what (is he saying)°?'
5 → IT8: =e:: 准备 è::: [siete preparati forse]
PRT prepare is are.2pl ready maybe
'= %e:: *zhunbei* is::: [are you ready maybe]%'

The data in Table 1 also suggest that the preferred strategy type for routine initiation was CRs, whereas CCs were more frequently used in subsequent reiterations. This sequence is very well exemplified in Excerpt 8. In line 61, IT3 produced a partial repetition of the trigger: in Section 3.5.1, this strategy will be interpreted as a CR for the elements of the word that were not grasped. After CH2's response, IT5 apparently understood what the intended word was, and repeated it for confirmation (line 65).

(8) Day4_#01

58	CH2:	[e	scrivania	è (0.2)	[桌-	写- =	
					[zh-	*x-*	
		[and	writing.desk	is	table	write	
		‘%and writing desk is%			[*zh- x-* =’		
59	IT3:				[°桌-		桌子°
					[°zhuō-	*zhuōzi°*	
					table-	table	
					‘[°tab-	table-°	
60	CH2:	=写字桌_					
		Xiězìzhuō					
		Write-character-table					
		‘=*Xiezizhuo_*’					
61 →	IT3:	写-					
		Xiě-					
		Write					
		‘*Xie-*’					
62	CH2:	scrive	è	[写 =			
				Xiě			
		write	is	write			
		‘%to write is%		[*xie* =’			
63	IT3:			[写			
				[xiě			
				write			
				‘[*Xie*’			
64	CH2:	= e	s- scri-	scrivania	写字桌_		
		= *e*	*s- scri-*	*scrivania*	*Xiězìzhuō_*		
		and		writing.table	write-character-table		
		‘= %and w- wr- writing table% *xiezizhuo_*’					
65 →	IT5:	°写字桌°					
		°*xiězìzhuō*°					
		write-character-table					
		‘°Writing table°’					

The overall distribution of each strategy type across the six days of observation is reported in Table 2 and Figure 1. As shown, CRs and CCs follow similar distributional patterns, and so do the other categories. In particular, the mean number and standard deviation of the CRs and CCs produced each day are very close, suggesting that the participants’ production of these two types of pragmatic strategies was relatively uniform. The data also show that the participants’ preference for CRs and CCs remained constant, with ISs and OSs only playing a marginal role.

Strategy type	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**M (SD)**
Clarification requests	4	15	11	10	8	18	11.00 (4.98)
Confirmation checks	1	13	9	15	7	16	10.17 (5.67)
Implicit signals	2	4	5	2	1	0	2.33 (1.86)
Other strategies	1	6	2	1	0	3	2.17 (2.14)
Total	8	38	27	28	16	37	25.67 (11.78)

Table 2. Strategy use across days of observation

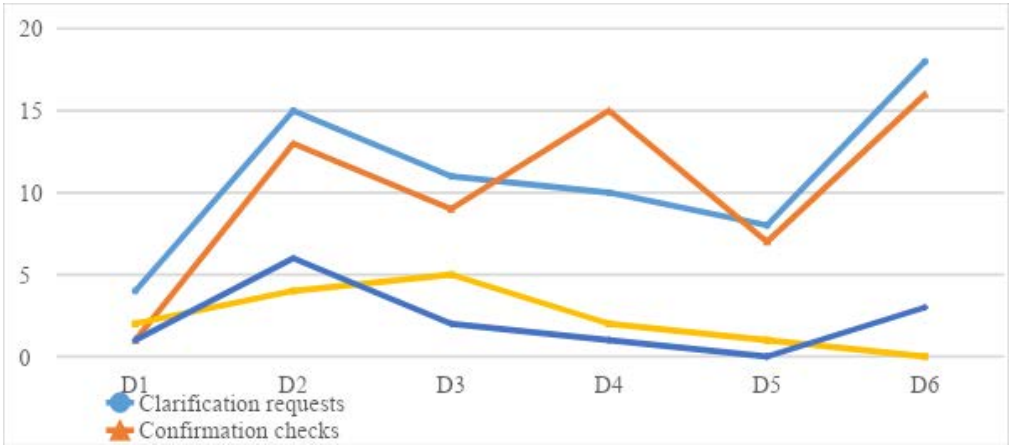


Figure 1. Strategy use across days of observation

Recall that in tandem language learning, each participant is a learner of their partners' L1. Therefore, code-switching is likely to occur, especially when non-understanding arises. As shown in Figure 2, this was indeed the case, with 45.45% (N = 70, M = 11.67, SD = 5.01) of the signals being uttered in Italian and 44.81% (N = 69, M = 11.50, SD = 6.44) in Chinese.

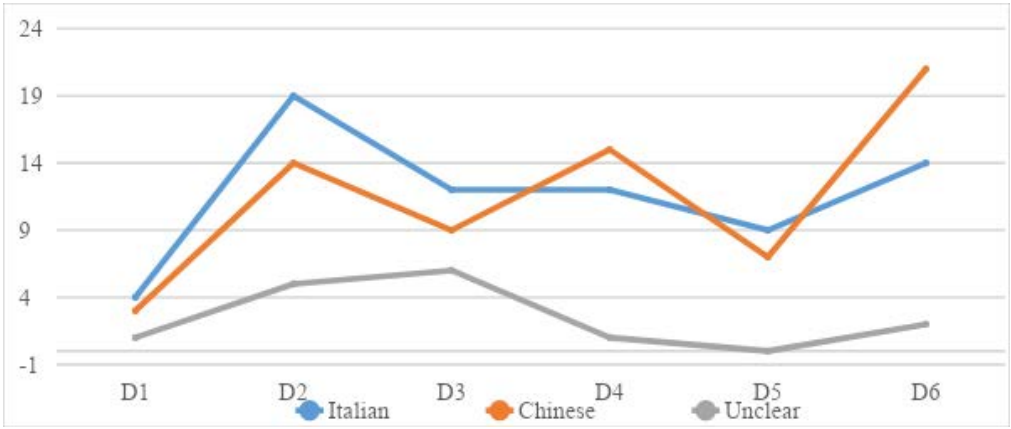


Figure 2. Use of Italian and Chinese

3.5 Clarification requests and confirmation checks

Considered the participants' preference for CRs and CCs, this Section will describe these two categories in more detail. In particular, it will focus on the form of the signal, that is, the linguistic and communicative resources the NNSs used to draw attention to a comprehension problem.

The data partially confirmed the realizations identified in previous studies (e.g., O'Rourke, 2005; Chen 2016). However, instances of Chinese-specific realizations also emerged.

3.5.1 Clarification requests

The CR forms observed in the data are reported in Table 3, together with an explanation and an example. The overall and day-by-day frequency of each form type is reported in Table 4.

Type	**Explanation**	**Example**
Explicit statement	The NNS explicitly indicates that non-understanding has occurred	*Méi tīngdǒng*没听懂 'I didn't understand'
Full repetition	The NNS repeats the entire trigger in a rising intonation, often with hesitations, mispronunciations etc.	*L- lei sing?*
Pragmatic marker	The non-understanding is signaled by means of a pragmatic marker, uttered with rising intonation	*Scusami?* 'Excuse me?'
Partial repetition	The NNS repeats part of the trigger (word, phrase, or utterance)	*Liú-* 流 'spread' (T: *liúxíng* 流行 'popular')
Question	The NNS asks explicit questions requiring clarifications	*Non è al contrario* 'isn't it the opposite'
Wh- pronoun	The NNS signals non-understanding using *wh-* pronouns in isolation	*Shénme*什么? 'What?'
'What is (T)?'	The NNS does not understand the trigger 'T' and explicitly asks what it is	*Xíngzhuàng* 形状*che cos'è?* 'What is *xingzhuang*?'
Others	• The NNS asks the other NNS for clarification	• *Spe che ha chiesto che* 'Wait what did they ask what'
	• Pragmatic marker + partial repetition	• *E::h rú-* 如-'U::h *ru-*?' (T: *rúguǒ* 如果 'if')

Table 3. CR forms

Type	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**Total**		
							N	**M (SD)**	**%**
Explicit statement	1	4	1	4	1	1	12	2.00 (1.55)	18.18
Full repetition	0	2	4	2	1	0	9	1.50 (1.52)	13.64
Pragmatic marker	2	2	0	0	2	2	8	1.33 (1.03)	12.12
Partial repetition	1	3	2	4	0	10	20	3.33 (3.56)	30.30
Question	0	0	0	0	3	4	7	1.17 (1.83)	10.61
Wh- pronoun	0	2	0	0	0	0	2	0.33 (0.81)	3.03
'What is (T)?'	0	1	4	0	1	0	6	1.00 (1.55)	9.09
Others	0	1	0	0	0	1	2	0.33 (0.52)	3.03

Table 4. Frequency of CR forms

Tables 4-5 show that the informants employed a variety of linguistics and communicative resources to signal non-understanding and request for clarification. The most frequent are partial repetition of the trigger and explicit statements of non-understanding, respectively. Other three relatively frequent forms are full repetitions, pragmatic markers (e.g., 'mh?', 'sorry?', 'uh?') and the use of the pattern 'what is T?', 'T' being the trigger and the question occurring both in Chinese ('T *shì shénme* 是什么?') and Italian ('T che cos'è?'). Explicit statements of non-understanding ('I didn't understand') also occurred in fixed forms, either Chinese ('*méi tīngdǒng* 没听懂') or Italian ('non ho capito').

Partial repetition of the trigger represents by far the most recurrent realization of CRs, its frequency almost doubling that of the next category, i.e., explicit statements. The repeated element can be a single morpheme in a compound, part of a phrase, or a segment of the entire utterance (Excerpts 9-11).

(9) Day4_#1

94	CH2:	…	questo (0.5)	鼠标垫.	
		…	*questo (0.5)*	*shǔbiāodiàn*	
			This	mouse.pad	
		'… %this% (0.5) mouse pad.'			
95 →	IT3:	dien?			
96	CH2:	垫 (.) 垫- 垫在下面.			
		diàn (.) diàn- diànzài xiàmian.			
		pad	pad	pad-at	underneath
		'Pad (.) pad- pad underneath'			

(10) Day6_#1

167	CH1:	>或者我等你的消息<=					
		>huòzhě	*wǒ*	*děng*	*nǐ*	*de*	*xiāoxi<*
		or	1sg	wait	2sg	GEN	news
		'>Or ((you can say)) I'll wait for your news<'					
((omitted))							
170 →	IT3:	你的?					
		nǐ	*de?*				
		2sg	GEN				
		'Your?'					
171	CH1:	消息 ((omitted))					
		xiāoxi ((omitted))					
		news					
		'News'					

(11) Day2_#1

5	CH3:	哦(.) 你们学汉语多久了?				
		ò (.)	*nǐmen*	*xué*	*hànyǔ*	*duōjiǔ-le?*
		PRT	2pl	study	Chinese	how-long-CRS?
		'How long have you been studying Chinese?'				
6 →	IT4:	e:: 你们?				
		e::	*nǐmen?*			
		PRT	2pl			
		'U::hm you?'				
7	CH3:	en:: 学汉语: (.) >学中文多久了 cinese<?				
		en::	*xué*	*hànyǔ: (.)*	*xué*	*zhōngwén duōjiǔ-le cinese*
		PRT	study	Chinese study	Chinese	how-long-CRS Chinese
		'En:: have studied Chine:se (.) >studied Chinese for how long %Chinese%<?'				

In Excerpt 9, the participants were engaged in a vocabulary task. In turns, they had to discuss the Italian or Chinese names of a series of objects in a picture. In the excerpt, negotiation was triggered by the word *shǔbiāodiàn* 鼠标垫 'mouse pad'. The component *shǔbiāo* 鼠标 'mouse' was already active in the previous context, as this word was discussed immediately before *shǔbiāodiàn* 鼠标垫. *Diàn* 垫 'pad', on the contrary, was completely novel: by repeating it, IT3 narrowed CH2's focus on the part of the word that required clarification. This example is opposite to Excerpt 8, line 61: in that case, IT3 repeated the only element that she comprehended (*xiě* 写 'to write'), and by doing so prompted CH2 to provide clarification to the novel ones.

Excerpt 10 is an example of a partially repeated phrase. Again, the repetition involved the elements which were familiar to the NNS (*nǐ de* 你的 [2sg-GEN] 'your'). As a confirmation that this strategy was indeed a CR and was interpreted as such by the interlocutors, in line 169 CH1 responded to the signal by supplying the segment of the trigger that IT3 failed to comprehend (*xiāoxi* 消息 'news').

Lastly, Excerpt 11 is an example of what Nakahama, Tyler, and van Lier (2001: 384-385) call a global trigger, that is, a trigger that involves the entire utterance rather than isolated lexical items or local morphosyntactic elements. In the excerpt, communication breakdown occurred as IT4 failed to understand CH3's question. Her partial repetition in line 6 suggests that she only understood the subject *nǐmen* 你们 'you'. CH3's subsequent response served as a clarification, as they provided a reformulation of their previous utterance, followed by the Italian translation of the words *hànyǔ* 汉语 and *zhōngwén* 中文 'Chinese language'.

In summary, in partial repetitions the repeated element is usually the one that the NNS understood, while the element that requires clarification is omitted (Excerpts 8, 10, and 11). However, the reverse is also possible (Excerpt 9). In this case, the repeated segment is the one that requires clarification.

As for code-switching, Figure 3 shows a preference for Chinese, which accounts for 57.58% of the occurrences (N = 38, M = 6.33, SD = 3.93). In comparison, Italian only accounts for 39.40% (N = 26, M = 4.33, SD = 1.86). However, if repetitions are excluded, the result is opposite, with a strong predominance of Italian (26 cases) over Chinese (9). Clearly, when adopting strategies other than repetitions, the participants are likely to code-switch to Italian.

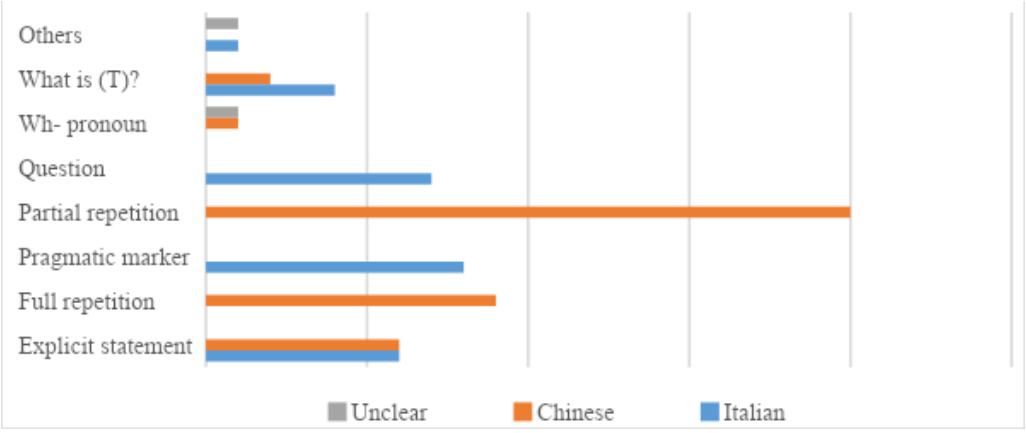


Figure 3. Code-switching in CRs

3.5.2 Comprehension checks

The complete list of the observed CC forms is reported in Table 5, whereas Table 6 shows the frequency of each form type.

Type	**Explanation**	**Example**
Confirmation question	The NNS explicitly asks for confirmation	*Nǐ bù xǐhuan* 你不喜欢? '(you mean) you don't like it?'
Full repetition	The NNS repeats the entire trigger	*Zuòfàn*做饭 'To cook'
Repetition + translation	The NNS repeats the trigger and provides a translation in Italian	*Dài* 带*come portare* '*Dai* as to bring'
Synonym	The NNS provides a Chinese synonym of the trigger	*Ah nánpéngyou* '男朋友 'Oh (the same as) boyfriend'
Translation	The NNS provides the Italian translation of the trigger, often with rising intonation	*Anello?* 'Ring?'
'(T) as in [compound]?'	The NNS asks if the trigger 'T' is part of a familiar compound word	*Kǎo (.) kǎoshì* 考 (.) 考试? '*Kao* (as in) *kaoshi* (do an exam)?'
Others	• Partial repetition	• *Dà* 大*?* 'Big?' (T: *dàmǐ* 大米 'rice')
	• Expansion	• *Ah era grande (.) dà de* 大的 (.) *hěn dà* 很大 'Oh (I get it) it was big… big… very big'
	• Word invention	• *Shǒujīdiàn* 手机垫? 'Mobile pad?'

Table 5. CC forms

Type	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**Total**		
							N	**M (SD)**	**%**
Confirmation question	0	2	0	1	1	0	4	0.67 (0.82)	6.56
Full repetition	1	4	0	4	4	5	18	3.00 (2.00)	29.51
Repetition + translation	0	0	1	1	0	6	8	1.33 (2.34)	13.11
Synonym	0	0	1	0	1	1	3	0.50 (0.55)	4.92
Translation	0	6	5	3	0	0	14	2.33 (2.73)	22.95
'(T) as in compound?'	0	0	0	5	1	4	10	1.67 (2.25)	16.39
Others	0	1	2	1	0	0	4	0.67 (0.82)	6.56

Table 6. Frequency of CC forms

Differently from CRs, CCs show two strongly predominant types, namely, full repetitions and translations. The frequency percentages of the two strategy types are very close; however, full repetitions are more uniformly distributed over the period of observation, while translations are more concentrated in the first four days.

An example of full repetition is that reported in Excerpt 2. Note that the NNS's inference is not necessarily accurate: in Excerpt 12, for instance, IT10 misinterpreted the NS's utterance *xǐ wǎn* 洗碗 'do the dishes' with the quasi-homophonous and assumedly more familiar *xīwàng* 希望 'to hope'.

(12) Day2_#5
115 IT9: … 你在家做饭 (.) 他做: (0.5) >什么<? =
… *nǐ* *zài* *jiā* *zuò-fàn (.)* *tā* *zuò:* *(0.5)*
… 2sg at home cook-rice 3sg.M do
>shénme<?
what
'… you cook at home (.) he do:es (0.5) >what<?'
116 CH4: [洗碗
[*xǐ* *wǎn*
wash bowl
'[he does the dishes'
((omitted))
120 → IT10: °°希望°°
Xīwàng
Hope
'°°To hope°°'

The other predominant CC type is translation. NNSs' translations can also be accompanied by pragmatic markers, as in Excerpt 13, or tag questions (e.g., *tipo hip-hop no?* 'like hip-hop, isn't it?'). These adjunct elements reinforce the illocutionary force of the utterance, suggesting that the NNS is trying to corroborate a hypothesis.

(13) Day3_#6
1 CH6: … 这个里面有 (.) 一百五十种 (.) 不同的: (.) 口味.
… *zhè* *ge* *lǐmiàn* *yǒu (.)* *yìbǎiwǔshí* *zhǒng (.)*
this CLF inside have one-hundred-fifty type
bùtóng *de (.)* *kǒuwèi.*
different DET taste
'… inside this ((place)) there are (.) one hundred and fifty types (.) of: different (.) flavours.'
2 → IT11: ah gusti.
PRT flavours
'Ah %flavours%.'

Interestingly, a relatively frequent strategy consisted in a combination of both repetition and translation. This hybrid form can also be followed by tag questions, as shown in Excerpt 14.

(14) Day6_#2
58 CH4: °很久°
°*hěn* *jiǔ*°
very long.time
'°A very long time°'

59 → IT4: ha 很旧- 旧 è: 旧 di: vecchio giusto?
ha hěn jiù- jiù è: jiù di: vecchio giusto?
PRT very old old is old of old right
'Ah, very old, %*jiu* i:s *jiu* li:ke old right%?'

60 CH4: e::n 很长的时间.
e::n hěn cháng-de shíjiān.
PRT vey long-DET time
'E::hm a very long time.'

A particularly interesting CC form consisted in proposing a familiar compound word or phrase which comprised the elements that caused the comprehension problem. In Excerpt 15, IT5 was trying to reconstruct the word *gǎnjué* 感觉 'to feel'. To do so, she suggested possible compounds or phrases in which the morphemes *gǎn* 感 'to feel' and *jué* 觉 'to sense' might occur.

(15) Day6_#1

87 CH1: 我感觉 (1.5) 我,
wǒ gǎnjué (1.5) wǒ,
1sg feel 1sg
'I feel (1.5) I,'

((omitted))

90 IT5: 感 [di:=
gǎn [di:=
feel of
'*Gan* [%as i:n%='

((omitted))

93 → IT5: =[感兴趣?
=[*gǎn xìngqu?*
feel interest
'=be interested?'

94 CH1: Sì (.) [quello 感.
Sì (.) [quello gǎn.
Yes that feel
'%Yes% (.) [%that% *gan*.'

95 IT3: [ok (1.0) 感-
[ok (1.0) *gǎn-*
ok feel
'[ok (1.0) *gan*-'

96 CH1: 觉.
jué
sense
'*jue*.'

97 → IT3: 觉得?
juéde.
sense-gain
'((As in)) to think?'

98 CH1: sì ((omitted))
yes
'%Yes% ((omitted))'

Lastly, a slight preference for Italian was observed (Figure 4). Italian CCs were 31 (50.82%, M = 5.17, SD = 3.49), while Chinese CCs were 27 (44.26). These data are confirmed if translation and repetitions – which are necessarily in Italian and Chinese, respectively – are not considered. In this case, CCs in Italian are 11, while CCs in Chinese are 8.

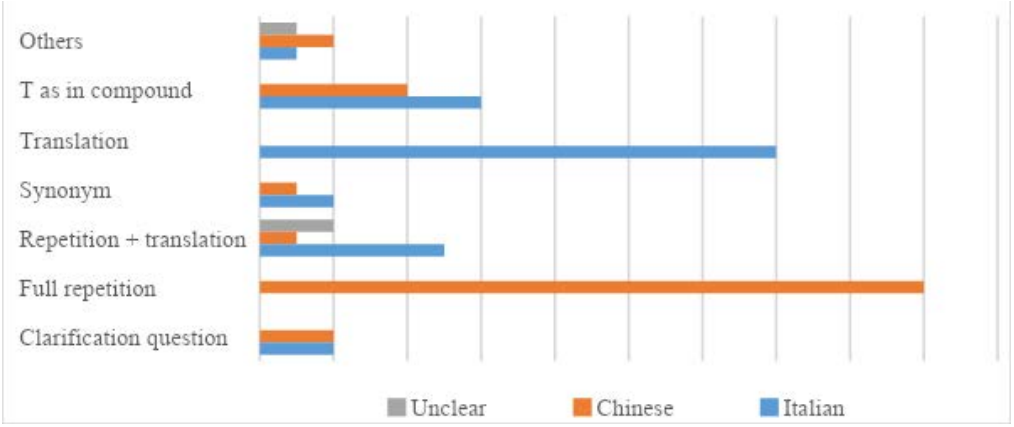


Figure 4. Code-switching in CCs

4. Discussion

The analysis of CFL learners' strategies to signal non-comprehension revealed a series of regularities, some confirming and some contradicting previous studies on NoM. Additionally, patterns emerged that are clearly target-language specific or related to the interactional context.

Firstly, the Italian participants recurred to both CRs and CCs. The frequency of occurrence of these two strategies is very close, suggesting that they were equally employed when dealing with comprehension problems (RQ1). This is in contradiction with Wang and Li (2015), who observed a predominance of CCs in NNS's signals. This discrepancy probably depends on classification criteria, particularly for what concerns repetitions. As reported in Section 3.3, repetitions are traditionally classified as CCs. Wang and Li (*ibidem*, 385) also adhered to this criterion, while in the present study repetitions were considered either CRs or CCs depending on speakers' co-constructed meaning as emerged from the interaction.

Despite their similar frequency, a difference between CRs and CCs can be observed if their sequential distribution in negotiation routines is considered. As shown, CRs tend to occur in routine-initiation position, whereas CCs are more likely to be used in routine reiterations, when the NS's response is not sufficient to solve the communication problem. It appears that miscommunications initially cause confusion to the NNSs, who choose to resort to CRs to explicitly signal incomprehension. After the NSs' response, the NNSs are able to make hypotheses on the interlocutor's intended meaning and use CCs to corroborate them. This pattern emerged very clearly from the data and provides an interesting starting point for future research on NoM.

Secondly, the participants employed a variety of linguistic resources to signal non-understanding, including pragmatic markers, *wh-* pronouns, reformulations, expansions, synonyms, etc. Yet, some resources – partial repetitions for CRs, full repetitions and translations for CCs – are unequivocally predominant (RQ2). These results resemble those of previous studies (e.g.,

O'Rourke 2005; Wang and Li 2015) and suggest that learners tend to adhere to a limited set of few, communicatively effective strategies.

Another point worth discussing is that concerning repetitions. As seen, repetitions reproduce the trigger either partially or in its entirety. In this regard, a clear difference was observed between repetitions serving as CRs and those serving as CCs. In CRs, the participants mostly used partial repetitions. The repeated part was the element of the word, phrase, or sentence that caused the incomprehension or, more typically, the only element that the NNS could grasp. In CCs, on the contrary, full repetitions were used almost exclusively. Note that full repetitions are also used as CRs, even though in this case they are simple phonetic imitations of the trigger. In other terms, they do not carry meaning. By contrast, when checking for confirmation the NNS has a clear hypothesis of what the problematic part might be. The repetition, therefore, is not aimed at requesting clarification; rather, it is aimed at receiving confirmation or rejection of such a hypothesis.

In sum, when asking for clarification, the NNSs tend to repeat the meaningful segment of the trigger, leaving the task of providing the opaque part to the NSs. On the contrary, when checking for confirmation, the NNSs tend to repeat the entire trigger, as they already have a clear representation of what the NSs' intended meaning might be.

Partial repetitions and the CC type that was labeled as "(T) as in compound" are highly Chinese-specific and show the NNS participants' awareness of Chinese word-building properties. Words in Modern Chinese can be analyzed as strings of monosyllabic morphemes (Norman 1988: 178). By repeating the meaningful morphemes of the trigger or suggesting possible compound words that contain the morpheme that requires disambiguation, learners demonstrate orientation towards the target language and its combinatory mechanisms. By contrast, strategies like translation are more L1-oriented, while other strategies (e.g., *wh-* questions) are not L1/L2 related. As such, they were produced both in Italian and Chinese, and even in English in one case ('what?').

This leads to another characteristic of the participants' signals, that is, the frequent code-switching and general preference for Italian (RQ3). Notably, the NSs also resorted to Italian in their responses to NNSs' signals, as shown in many of the excerpts in Section 3. Similar findings also emerged from O'Rourke's (2005) analysis of Irish and German NoM in computer-mediated tandem. In that study, English was the dominant language, and four out of five cases of explicit suggestions for repair consisted of German learners' call for translations into English. Following O'Rourke, it can be concluded that Italian assumed the status of a lingua franca, a safe solution to most communication problems. This might be due to the proficiency imbalance between the two groups of learners, even though they assumedly had a similar L2 level. Another reason may depend on the fact that Italian was the dominant language in the learning environment, as all the interactions took place in an Italian university. Either way, without drawing hasty conclusions on the amount of L2 acquisition, it is reasonable to assume that Chinese learners presumably benefited more from the volume of L2 input and output, thus altering the linguistic, pedagogical, and affective nature of a tandem exchange (*ibidem*, 458).

As a last remark, learners' signal strategies, though showing a clear perlocutionary effect, were particularly error-prone or syntactically incomplete. An example is shown in Excerpt 15, line 97: a more accurate way to ask if a morpheme is part of a compound would be using the "compound – DET – morpheme" construction, in that case *juéde de jué* 觉得的觉 [think – DET – sense] '*jue* as in *juede* [to think]'. IT3, however, only conveyed the compound (*juéde* 觉得? 'to think?'), whereas the other elements of the construction were omitted. CH1's acknowledgment in line 98 suggests that they understood IT3's intention; still, the illocutionary force of IT3's

utterance was not explicit, and the burden of its interpretation was entirely on the NS's part. Following Taguchi and Roever (2017: 120), it appears that low-ability learners seem to rely on their interlocutors' cooperation and use them as a "resource to accomplish social actions by proxy". This strategy does not necessarily lead to pragmatic failure, but it highlights NNSs' overreliance on the NSs' ability to simplify and facilitate the interaction. Such inaccuracies were never addressed by the NSs. This is in line with previous research on corrective feedback in tandem learning, during which "students are more concerned with communication than with correction. In other words, structural features are treated mainly when incorrect usage causes a problem in communication" (Cappellini 2016: 14; see also Cappellini and Zhang 2013). In fact, one of the drawbacks of tandem is its tolerance to errors, particularly syntactical ones (Little 2003). This diminished the chances for the Italian participants to receive negative feedback and become more efficient negotiators.

5. Conclusions

This study analyzed Italian and Chinese learners' interactions in a tandem learning context and classified Italian CFL learners' pragmatic strategies to signal non-understanding. The data showed the participants' tendency to use a limited set of linguistic resources, mostly consisting of full or partial repetitions of the trigger, as well as a general preference for Italian, used as a lingua franca to interrupt the flow of the conversation and focus on the problem-solving activity. Empirical evidence also suggested a sequential preference for CRs and CCs, which tend to occur in different positions of the negotiation routine (initiation and reiterations, respectively).

According to Varonis and Gass (1985), NoM is more frequent among NNS-NNS than NS-NNS pairs, as NNSs' "shared incompetence" (*ibidem*, 84) allows them to put the conversation on hold without fear of losing face. The results of the present study suggest a more encouraging scenario. Indeed, the participants engaged frequently and very deeply into NoM, and even though it is impossible to ascertain whether the Italian learners signalled each and every instance of non-comprehension, cases of abandoned negotiations were not attested. Assumedly, in a tandem context "partners' awareness of the learning purpose of the exchange and their shared status as learners, leads them to use direct failure signals that would in nonpedagogical situations be regarded as face-threatening" (O'Rourke 2005: 449). Therefore, tandem exchanges are likely to encourage NoM, even though little attention is paid to the linguistic accuracy of the signals themselves.

These results may be particularly relevant in terms of pedagogical implications. According to McKay (2010: 239), negotiation practices should be introduced in the language classrooms. In order to increase linguistic accuracy and enhance communicative efficiency, learners should be given samples of authentic interactions and then asked to identify those strategies that help the meaning-construction process (De Bartolo 2014: 460). In CFL, Wang and Li (2015: 389) suggest focusing learners' attention on the fixed or semi-fixed chunks that NSs use as signals. Regular patterns for signaling non-understanding are reported in Jin (2011: 86). According to the author, typical Chinese CRs consist of questions (… *shì shénme yìsi* 是什么意思 'what does … mean?'), imperatives (*qǐng nǐ màn yìdiǎn* 请你慢一点 'a bit slower, please'), or a combination of statements and imperatives (*wǒ méi tīng dǒng, qǐng zài shuō yí cì* 我没听懂，请再说一次 'I didn't understand, please repeat'); as for CCs, frequently recurring formulae are *nǐ de yìsi shì bu shì* 你的意思是不是… or *nǐ shì bu shì shuō* 你是不是说… 'Do you mean…'. These and other forms may be introduced to and practiced by learners as effective strategies to overcome communication problems, while at the same time sounding more nativelike.

Due to the small sample size, the results of this study cannot be used to draw general conclusions on CFL learners' practices during NoM. Still, they provide valuable insight into such practices, hopefully inspiring future research on negotiation in CFL/CSL.

References

Bazzanella, Carla, and Rossana Damiano. 1997. "Il fraintendimento nelle interazioni quotidiane: proposte di classificazione." *Lingua e stile* 32 (3): 173-200.

Bazzanella, Carla, and Rossana Damiano. 1998. "The Interactional Handling of Misunderstanding in Everyday Conversations." *Journal of Pragmatics* 31: 817-836.

Brammerts, Helmut. 2003. "Apprendimento autonomo in tandem: sviluppo di un modello." In *Apprendimento autonomo delle lingue in tandem*, ed. by Georg Hehmann, and Donatella Ponti, 13-20. Torino: Trauben.

Cappellini, Marco. 2016. "Roles and Scaffolding in Teletandem Interactions: A Study of the Relations between the Sociocultural and the Language Learning Dimensions in a French-Chinese Teletandem." *Innovation in Language Learning and Teaching* 10 (1): 6-20. doi: 10.1080/17501229.2016.1134859.

Cappellini, Marco, and Meng Zhang. 2013. "Étude des négociations du sens dans un tandem par visio-conférence." *Recherches en didactique des langues et des cultures* 10 (2): 1-26. doi: 10.4000/rdlc.1576.

Chen, Julian C. 2016. "EFL Learners' Strategy Use during Task-based Interaction in Second Life." *Australasian Journal of Educational Technology* 32 (3): 1-17. doi: 10.14742/ajet.2306.

Comeau, Liane, Fred Genesee, and Morton Mendelson. 2010. "A Comparison of Bilingual Children's Conversational Repairs." *First Language* 30 (3-4): 354-374. doi: 10.1177/0142723710370530.

De Bartolo, Anna Maria. 2014. "Pragmatic Strategies and Negotiation of Meaning in ELF Talk." *EL.LE* 3 (3): 453-464. doi: 10.14277/2280-6792/115p.

Doughty, Catherine. 2000. "Negotiating the L2 Linguistic Environment." *University of Hawai'i Working Papers in ESL* 18 (2): 47-83. <http://hdl.handle.net/10125/40806> (06/2021).

Fernández-García, Marisol, and Asunción Martínez-Arbelaiz. 2002. "Negotiation of Meaning in Nonnative Speaker-Nonnative Speaker Synchronous Discussion." *CALICO Journal* 19 (2): 279-294. doi: 10.1558/cj.v19i2.279-284.

Gass, Susan M., and Evangeline M. Varonis. 1994. "Input, Interaction, and Second Language Production." *Studies in Second Language Acquisition* 16 (3): 283-302. doi: 10.1017/S0272263100013097.

Gass, Susan M., Alison Mackey, and Lauren Ross-Feldman. 2005. "Task-based Interactions in Classroom and Laboratory Settings." *Language Learning* 55 (4): 575-611. doi: 10.1111/j.0023-8333.2005.00318.x.

Hepburn, Alexa, and Galina B. Bolden. 2013. "The Conversation Analytic Approach to Transcription." In *The Handbook of Conversation Analysis*, ed. by Jack Sidnell, and Tanya Stivers, 57-76. Chichester-Malden, MA: Wiley-Blackwell. doi: 10.1002/9781118325001.ch4.

Jefferson, Gail. 2004. "Glossary of Transcript Symbols with an Introduction." In *Conversation Analysis: Studies from the First Generation*, ed. by Gene H. Lerner, 13-31. Amsterdam: John Benjamins. doi: 10.1075/pbns.125.02jef.

Jin, Honggang 靳洪刚. 2011. "Xiandai Hanyu jiaoxue de shi da yuanze 现代语言教学的十大原则 [Ten Major Principles of Modern Language Teaching Methodology]." *Shijie Hanyu Jiaoxue* 世界汉语教学 25 (1): 78-98. doi: 10.13724/j.cnki.ctiw.2011.01.009.

Lee, Helen, Regine Hampel, and Agnes Kukulska-Hulme. 2019. "Gesture in Speaking Tasks beyond the Classroom: An Exploration of the Multimodal Negotiation of Meaning via Skype Videoconferencing on Mobile Devices." *System* 81: 26-38. doi: 10.1016/j.system.2018.12.013.

Lee, Jee W., Hongyin Tao, and Ping Lu. 2017. "Transcribing Mandarin Chinese Conversation: Linguistic and Prosodic Issues." *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology* 7 (5): 787-799. doi: 10.14257/ajmahs.2017.05.70.

Lee, Lina. 2001. "Online Interaction: Negotiation of Meaning and Strategies Used among Learners of Spanish." *ReCALL* 13 (2): 232-244. doi: 10.1017/S0958344001000829a.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkley, CA-Los Angeles, CA-London: University of California Press.
Little, David. 2003. "Imparare le lingue in tandem e autonomia dell'apprendente." In *Apprendimento autonomo delle lingue in tandem*, ed. by Georg Hehmann, and Donatella Ponti, 21-27. Torino: Trauben.
Long, Michael H. 1980. *Input, Interaction, and Second Language Acquisition* (unpublished doctoral dissertation). Los Angeles, CA: University of California.
Long, Michael H. 1983. "Native Speaker/Non-Native Speaker Conversation and The Negotiation of Comprehensible Input." *Applied Linguistics* 4 (2): 126-141. doi: 10.1093/applin/4.2.126.
Long, Michael H. 1985. "Input and Second Language Acquisition Theory." In *Input in Second Language Acquisition*, ed. by Susan M. Gass, and Carolyn G. Madden, 377-393. Rowley, MA: Newbury House Publishers.
Long, Michael H. 1996. "The Role of the Linguistic Environment in Second Language Acquisition." In *Handbook of Second Language Acquisition*, ed. by William C. Ritchie, and Tej K. Bathia, 413-468. New York, NY: Academic Press.
Lu, Ping 陆萍, Jee W. Lee 李知沅, and Hongyin Tao 陶红印. 2014. "Xiandai Hanyu kouyu zhong teshu huayu yuyin chengfen de zhuanxie yanjiu 现代汉语口语中特殊话语语音成分的转写研究 [Discourse Properties of Some Special Sound Elements and Their Transcription Treatment]." *Yuyan Kexue*语言科学 13 (2): 113-130. doi: 10.7509/j.linsci.201311.028418.
Lu, Yang. 2017. "Exploring the Criterion-validity of HSK Levels 3 and 4: Are Assessments and CEFR Standards Related?" In *Teaching and Learning Chinese in Higher Education: Theoretical and Practical Issues*, ed. by Yang Lu, 35-56. Abingdon: Routledge.
Masrizal, Mahmud. 2014. "The Role of Negotiation of Meaning in L2 Interactions: An Analysis from the Perspective of Long's Interaction Hypothesis." *Studies in English Language and Education* 1 (2): 96-105. doi: 10.24815/siele.v1i2.1829.
McKay, Sandra L. 2010. "Pragmatics and EIL Pedagogy." In *English as an International Language: Perspectives and Pedagogical Issues*, ed. by Farzad Sharifian, 227-241. Bristol: Multilingual Matters. doi: 10.21832/9781847691231-015.
Nakahama, Yuko, Andrea Tyler, and Leo van Lier. 2001. "Negotiation of Meaning in Conversational and Information Gap Activities: A Comparative Discourse Analysis." *TESOL Quarterly* 35 (3): 377-405. doi: 10.2307/3588028.
Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge UP.
Oliver, Rhonda. 2002. "The Patterns of Negotiation for Meaning in Child Interactions." *The Modern Language Journal* 86 (1): 97-111. doi: 10.1111/1540-4781.00138.
O'Rourke, Breffni. 2005. "Form-focused Interaction in Online Tandem Learning." *CALICO Journal* 22 (3): 433-466. doi: 10.1558/CJ.V22I3.433-466.
Pica, Teresa. 1994. "Research on Negotiation: What Does it Reveal about Second-language Learning Conditions, Processes, and Outcomes?" *Language Learning* 44 (3): 493-527. doi: 10.1111/j.1467-1770.1994.tb01115.x.
Pica, Teresa, and Catherine Doughty. 1985. "Input and Interaction in The Communicative Language Classroom: A Comparison of Teacher-fronted and Group Activities." In *Input in Second Language Acquisition*, ed. by Susan M. Gass, and Carolyn G. Madden, 115-135. Rowley, MA: Newbury House.
Pica, Teresa, Lloyd Holliday, Nora Lewis, Dom Berducci, and Jeanne Newman. 1991. "Language Learning through Interaction: What Role Does Gender Play?" *Studies in Second Language Acquisition* 13 (3): 343-376. doi: 10.1017/S0272263100010020.
Qian, Daojing 钱道静. 2010. "Duiwai Hanyu ketang jiaoxue zhong de yuyi xieshang jiaoji huodong 对外汉语课堂教学中的语义协商交际活动 [Negotiation of Meaning in Chinese Classroom Interaction]." *Dajia*大家 12: 128-129.
Sabatini, Francesco. 1985. "L'italiano dell'uso medio': una realtà tra le varietà linguistiche italiane." In *Gesprochenes Italienisch in Geschichte und Gegenwart*, ed. by Günter Holtus, and Edgar Radtke, 154-184. Tübingen: Gunter Narr Verlag.
Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. "A Simplest Systematics for the Organization of Turn-taking for Conversation." *Language* 50 (4): 696-735.

Su, Er 苏尔, and Xiaoqing Hu 胡晓清. 2017. "Di er yuyan xide zhong de 'yiyi xieshang' yanjiu suping 第二语言习得中的'意义协商'研究述评 [A Review of Studies on 'Negotiation of Meaning' in Second Language Acquisition]." *Ludong Daxue Xuebao (Zhexue Shehui Kexue Ban)*鲁东大学学报（哲学社会科学版）34 (4): 71-78.

Swain, Merrill. 1985. "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development." In *Input in Second Language Acquisition*, ed. by Susan M. Gass, and Carolyn G. Madden, 235-253. Rowley, MA: Newbury House Publishers.

Swain, Merrill. 2000. "The Output Hypothesis and Beyond: Mediating Acquisition through Collaborative Dialogue." In *Sociocultural Theory and Second Language Learning*, ed. by James P. Lantolf, 97-114. Oxford: Oxford UP.

Taguchi, Naoko, and Carsten Roever. 2017. *Second Language Pragmatics*. Oxford: Oxford UP.

Van der Zwaard, Rose, and Anne Bannink. 2019. "Negotiation of Meaning in Digital L2 Learning Interaction: Task Design versus Task Performance." *TESOL Quarterly* 54 (1): 56-89. doi: 10.1002/tesq.537.

Varonis, Evangeline M., and Susan M. Gass. 1985. "Non-native/non-native Conversations: A Model for Negotiation of Meaning." *Applied Linguistics* 6 (1): 71-90. doi:10.1093/applin/6.1.71.

Wang, Pingli 王萍丽, and Yanli Li 李彦霖. 2015. "Yuyi xieshang de xiaoyong yu huabu goucheng – jiyu muyuzhe he feimuyuzhe ziran yuyan hudong de ge'an yanjiu 语义协商的效用与话步构成——基于母语者和非母语者自然语言互动的个案研究 [Effects of Negotiation of Meaning and Turn-taking Structure: A Case Study of Naturalistic Interactions between a Native Speaker and a Non-native Speaker]." *Shijie Hanyu Jiaoxue* 世界汉语教学 29 (3): 377-392.

Wang, Yuping. 2006. "Negotiation of Meaning in Desktop Videoconferencing-supported Distance Language Learning." *ReCALL* 18 (1): 122-145. doi: 10.1017/S0958344006000814.

Wu, Ruey-Jiuan R. 2016. "Doing Conversation Analysis in Mandarin Chinese. Basic Methods." *Chinese Language and Discourse* 7 (2): 179-209. doi: 10.1075/cld.7.2.01wu.

Zhao, Lei 赵雷. 2015. "Renwuxing kouyu ketang Hanyu xuexizhe xieyi hudong yanjiu 任务型口语课堂汉语学习者协商互动研究 [Research on Negotiated Interactions among Chinese Learners in Task-based Spoken Classroom]." *Shijie Hanyu Jiaoxue*世界汉语教学 29 (3): 362-376.

Comunicazione

Communication

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

OPEN ACCESS

Citation: S.C. Crocchi, L. Dal Pozzo (2021) Framing Covid-19: A comparative analysis between institutional and press communication in Finland and Italy in 2020. *Qulso* 7: pp. 317-343. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12013

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Framing Covid-19: A comparative analysis between institutional and press communication in Finland and Italy in 2020

Serena Camilla Crocchi, Lena Dal Pozzo
Università degli Studi di Firenze
(<serenacamilla.crocchi@unifi.it>; <lena.dalpozzo@unifi.it>)

Abstract:

In the present work we focus on institutional communication in a period of crisis. In particular we are interested in the communications of the Italian and Finnish governments during the health emergency raised by the COVID-19 pandemic. We analyse and compare two press releases given by the Italian and Finnish Prime Ministers at the very beginning and at the end of the so-called first stage of the pandemic. The analysis aims at i) describing the effectiveness of the speeches applying six parameters that we previously identified as crucial for efficient communication, and ii) exploring the lexical choice of the Prime Ministers in relation to it.

Keywords: *COVID-19, Crisis Communication, Infodemic, Institutional Communication, Storytelling*

1. Framing the study

The consequences triggered by the spread of the pandemic, unleashed following the spread of the SARS-CoV-2 virus, are still incalculable, after a year and a half from the first confirmed case in China. The initial impossibility and the general difficulty of stemming the spread of the disease in a highly globalized world have put the economy of most countries in crisis, without distinction. The pandemic has highlighted, in many cases, the weaknesses of government and health systems, it has shaken the foundations of corporate organizational systems and has redesigned the ways in which knowledge and information are used and provided. The spread of the epidemic and the attempts to remedy it have forced humanity, in essence, to question everything. No sector has therefore remained unscathed, immobile or equal to itself in recent months. The areas most affected

by this paradigmatic revolution are, above all, medicine, science and politics, but the element that stands out most strongly in this case analysis, as a whole, and for the aim of this paper, is undoubtedly the communicative one. As in crisis situations in general, also in the COVID-19 pandemic communication has gained utmost importance. The kind of communication involved by the COVID-19 emergency can be considered crisis communication along the lines of Reynolds and Seeger (2005) who discuss the communication types during crisis and identify the crucial differences between crisis and emergency communication. In crisis communication due to a health emergency, health authorities generally track the origin of the disease, undertake actions to stop the contamination, issue warnings and recalls, and provide the public with information about the symptoms, treatments and ways to avoid exposure (Fischhoff and Downs 2001; Ulmer and Sellnow 2000 as in Reynolds and Seeger 2005: 44).

The pervasive use and massive success of social media, starting from early 2000s, have led to the redefinition of the power relationships within communication exchanges and, despite the many benefits, have fuelled some trends and generated real informative disorders. In this article we focus on institutional and press communication in Italy and in Finland, and we leave aside the communication spread through social media and the related networks. It is well known that solid leadership, empathy, and alignment of science and politics are crucial factors for clear communication. Even though Finland and Italy can be considered as equally good representatives of good communication in the pandemic scenario with their frequent, truthful and clear institutional communications, we aim at describing here how the language of these communications may vary in terms of lexical choice and discourse structure and how and in which measure clearness, truthfulness, coherency, and empathy emerge. For this purpose, we focus on two talks of the prime ministers and on the day after communications in the main newspapers in two different moments in Italy and in Finland. It is well-known in current communication literature (Albu and Flyverbom 2019; Menon and Goh 2005) that effective institutional communication during a crisis necessarily involves a variety of elements. Two of the most common fallacies are communicating too less or too much to the selected audience (Casalegno and Civera 2016).

For the purpose of this study, we take into consideration the following parameters (adapted from Casalegno, *et al.* 2020) as crucial elements for efficient institutional communication:

1. Harm-reducing, honest and accurate information;
2. Simple and quick message;
3. Clear objectives and goals and how they are presented;
4. Provision of proof of tangible actions;
5. Coherency among institutional communications and media (press);
6. Communication strategies.

In this paper we first present how the COVID-19 emergency has developed in Italy and in Finland and then we aim at investigating how the above mentioned six parameters emerge in four official communications given by the Italian (2) and the Finnish (2) governments in two different moments of the COVID-19 health emergency. In particular, we focus on the lexical choice and syntax of the discourses and report on what kind of differences, if any, can be observed in the communications of Italian and Finnish prime ministers. Finally, we observe the COVID-19-related headings in two national daily newspapers in Italy and in Finland (*Corriere della Sera* and *Helsingin Sanomat*, respectively) in order to check how many headings are on the governmental communication of the day before and if and how lexical choice and communicative strategies differ.

The paper is organized as follows: section 2 resumes the beginning of the pandemic in Italy and in Finland. In section 3 different types of challenges faced by institutional communication and information channels are presented. Section 4 starts the second part of the paper devoted to the talks of the Prime Ministers and to the day-after newspapers. In Section 5 the results are discussed from a comparative point of view and Section 6 concludes the paper.

2. The story so far

On February 11th 2020 the World Health Organization (WHO from now on) announced that the respiratory disease caused by the new coronavirus was named COVID-19. On January 30th, 2020 the Istituto Superiore di Sanità (ISS from now on) confirmed the first two cases of COVID-19 infection in Italy and on February 21st, it confirmed the first autochthonous case in Italy. Finally, on 11th March 2020, the WHO declared COVID-19 to officially be a pandemic after 3 months when China broke out the first outbreak of coronavirus disease in December 2019 in Wuhan city[1] (population 11 million). When the Chinese government quarantined the city of Wuhan in mid-January 2020 to prevent the spread of COVID-19, authorities in Europe had not imagined that the coronavirus would threaten them so invasively. While the first confirmed case of COVID-19 in Italy was on January 30th 2020, and additional cases of COVID-19 infection emerged in Codogno (Milan), the Italian government imposed a quarantine a month later on February 21th, 2020 for infected people and their contacts and for those who came back from China. Subsequently tighter measures were imposed for northern Italy, with the declaration of the so-called Red Zones on March 8 "*Decreto Zone Rosse*" 'Decree on Red Zones' and with the "*Decreto #iorestoacasa*" '#stayathome decree' on March 11th, 2020 imposing a total lockdown of public life. Moreover, from March 22nd "*Decreto Chiudi Italia*" 'Close Italy Decree', it was forbidden to go in public parks, playground areas, and to play outside. Sports activities were allowed only near residences. In order to counter the spread of the pandemic and monitor its progress, the app *Immuni* was launched in Italy from June 2020. The experimentation started in four regions at first and was then extended to the whole national territory. It has been quite contested and the number of downloads has been very low. At the beginning of the second wave in Italy, in autumn 2020, Giuseppe Conte and his staff did not mention the app once as part of their strategy confirming that it has been a failure. At the time of this writing, about eighteen months have passed since the first confirmed case in Italy and there have been reports of over 182.209.207 COVID-19 cases worldwide and 3.945.829 million deaths globally, and 166.772.587 recovered.[2] Specifically in Italy (population[3] 59.258.000), there are 4.258.456 COVID-19 cases, 127.500 deaths and 4.076.274 recovered. At the time of this writing Italy is experiencing a *White Zones* phase in which movements are allowed and restrictions are very limited and people are not wearing masks outdoors.

In Finland (5,548,361 population)[4], there are 95.168 COVID-19 cases, 969 deaths and 46.000 recovered.[5] The first suspicions case that the virus had arrived in Finland were aroused on 24 January when a Chinese tourist visiting Inari, Lapland, was brought to the Ivalo Health Centre suffering from flu-like symptoms but the first coronavirus case in Finland was diagnosed

[1] <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (06/2021).
[2] <https://www.worldometers.info/coronavirus/> (06/2021).
[3] <https://www.istat.it/it/archivio/257243> (06/2021).
[4] <https://www.populationpyramid.net/it/finlandia/2021/> (06/2021).
[5] <https://www.worldometers.info/coronavirus/country/finland/> (06/2021).

for a Chinese tourist in a holiday resort in Lapland on January 29th 2020. On February 13th 2020 COVID-19 infection was included in the list of generally hazardous communicable diseases by amending the Government Decree on Communicable Diseases. However, the situation with the spread of COVID-19 in Finland started to evolve only in early March, boosted by travels to the Alps during the winter holidays in February. COVID-19 epidemic was considered to have started in Finland in mid-March soon after the WHO had announced the COVID-19 outbreak as a pandemic. On March 16th 2020 the Finnish Government announced a state of emergency due to the coronavirus outbreak and consequently it implemented several strategies in order to improve social and physical distancing which aimed at slowing the spread and protecting especially groups at risk. So far, Finland has been successful in restraining the pandemic. Compared to many countries, the pandemic landed in Finland relatively late. This fact allowed early implementation of restrictive regulations and recommendations. However, these measures which have been obeyed well by the population have probably had several negative, unintended consequences (Tiirinki *et al*. 2021) defined as changes brought by an intervention other than those it aims to achieve (Jabeen 2016). They can be far-reaching and just as important as the intended consequences and they can affect all groups of the population as well as all sectors of society (e.g. health, education, environment, economy, law). As stated in the study conducted by Turcotte, Gali and Ridde (2021) there is scientific evidence suggesting that COVID-19 mitigation measures can trigger a wide range of desirable and undesirable unintended consequences within and outside of healthcare systems.

The Finnish Institute for Health and Welfare (*Terveyden ja hyvinvoinnin laitos*, THL from now on) is the main national statistical and registry authority for the health and social sector, and it is also responsible for collecting the COVID-19 data. The COVID-19 data are released at the THL website. The collected data have been published daily from the start of the pandemic. However, THL has been criticized for not giving out detailed information on the deaths and patients treated in the hospitals. For instance, the breakdown of the cases according to age and sex was not initially reported. Detailed data on intensive care were only released weekly by the IC Coordinating Office. THL has also been criticized for not releasing the methodological details and assumptions used for modelling the epidemic.[6] There is also no published data on contact tracing as there is no tracing in the successively developed app *koronavilkku* (downloadable since 31.8.2020) through which only the information that one could have been infected in the previous 14 days is provided. This is said to protect sensible information and the privacy of citizens.

2.1 Phases of the Covid-19 Pandemic and Measures of containment in Italy

In Italy as the health system is regionally based, COVID-19 emergency actions are based on a general lockdown imposed by national authority and then management at local level by 21 regional authorities. Therefore, the pandemic response plan developed by each region led to different approaches. It is possible to identify three main waves since the beginning of the pandemic in Italy.

[6] This seems related to a general willingness to avoid any kind of *over-communication* (or *infodemic* as formulated in Zaracostas 2020) that can create fear and confusion instead of reassuring the population and thus be self-defeating for the purposes of the government.

First wave - *Lockdown and 3-phase split*:

Phase 1: Lockdown (March 9 to May 3 2020).
Phase 2: Relaxation of containment measures (May 4 - June 14 2020).
Phase 3: Coexistence with COVID-19 (June 15 - October 7 2020).

Second wave - *Resumption of the infection*:

8 October - 5 November: new restrictive actions.
From 6th November: containment for different scenarios (Curfew and establishment of yellow, orange and red zones), Measures for the holiday season (December 21 - January 15 2021), Establishment of white zones (from January 16 2021).

Third wave - *The spread of variants*:

Tightening of containment measures (6 March - 25 April 2021).
Relaxation of containment measures (from 26th April 2021).
Re-openings and Green Certification.

The first lockdown began around 21st February 2020, covering ten municipalities of the province of Lodi in Lombardy and one in the province of Padua in Veneto, and affecting around 50.000 people. The lockdown was initially meant to last until March 6th. Citizens were permitted to leave their homes to look for supplies such as food and medicine, to go to workplaces or to schools was not allowed, and public gatherings were prohibited. Train services also bypassed the region. Early on Sunday 8th March 2020, Italian Prime Minister Giuseppe Conte announced the expansion of the quarantine zone to cover much of northern Italy, affecting over sixteen million people, restricting travel from, to or within the affected areas, banning funerals and cultural events, and requiring people to keep at least one metre of distance from one another in public locations such as restaurants, churches and supermarkets. Conte later clarified in a press conference that the decree was not an "absolute ban", and that people would still be able to use trains and planes to and from the region for proven work needs, emergencies, or health reasons. Additionally, tourists from outside were still permitted to leave the area. Restaurants and cafes were permitted to open, but operations were limited to between 6:00 and 18:00, while many other public locations such as gyms, nightclubs, museums and swimming pools were closed altogether. Businesses were ordered to implement "remote working" to permit their employees to work from home. The decree, that was into force until 3rd April 2020, additionally cancelled any leave for medical workers, and allowed the government to impose fines or up to three months' jail for people caught leaving or entering the affected zone without permission. The decree also implemented restrictions on public gatherings elsewhere across Italy. With this decree, the initial "Red Zone" was also abolished (though the municipalities were still within the quarantined area). The lockdown measures implemented by Italy were considered the most radical measures implemented against the outbreak outside of the lockdown measures implemented in China. At the time of the decree, over 5.800 cases of coronavirus had been confirmed in Italy and 233 deaths. A draft of the decree had been leaked to the media late on Saturday night before it went into effect and was published by the newspaper *Corriere della Sera*, resulting in panic within the to-be-quarantined areas and prompting reactions from politicians in the region. The newspaper *La Repubblica* reported that hundreds of people in Milan rushed

out to leave the city on the last trains on Saturday night, as a part of a general rush to leave the new expanded red zone. *La Repubblica* later reported that this was an exaggeration and that, through an analysis of telephone cells, less than 1.000 people had left Milan for the Southern regions on 7th March; by comparison, on 23rd February, about 9.000 people left Milan towards South Italy. However, within hours of the decree being signed, media outlets reported that relatively little had changed, with trains and planes still operating to and from the region, and restaurants and cafes operating normally. According to the Ministerial Decree of 26th April 2020 the masks had become mandatory in confined spaces or outdoors when the possibility of maintaining the necessary physical distance is not possible or guaranteed.

2.2 Phases of the Covid-19 Pandemic and Measures of containment in Finland

In Finland the action plan for implementing the Government's hybrid strategy divides the COVID-19 epidemic into three phases. This classification is used to assess the development of the epidemic and the need for recommendations and restrictions and to ensure that they are appropriately targeted in the state.

1. *Perustaso "baseline" summer 2020 (June-beginnig of September)*. The epidemic is at stable level, and incidence is low. Local and regional transmission chains occur only occasionally. Contamination chains are manageable, and the people exposed can be traced without delay. New cases are either random isolated cases or most of them are detected among people in quarantine.

2. *Kiihtymistaso* "increasing stage" (traceability of the contaminations, hospital situation under control). The epidemic is accelerating, and regional incidence is higher than at the stable level. This may mean that the percentage of people testing positive is above one and that large-scale exposures occur.

Contact tracing manages to identify a significant proportion of the sources of infection.

There are many local and regional transmission chains.

3. *Leviämistaso* "diffusion stage" (20-25 cases in 7 days; less than half of the contaminations can be traced; problems in the hospitals). The epidemic continues to accelerate. This may mean that the percentage of people testing positive is above two and that the need for inpatient and intensive care is projected to increase. Cases are spreading at the regional level or more widely through the population. Contact tracing is becoming more difficult.[7]

In Finland the first press release given by the government on the COVID-19 epidemic dates to 27th February 2020. This very first communication the Prime Minister Sanna Marin and the Council of Ministers want to reassure the population on the diffusion of COVID-19 in Finland. It is said that Finland is not involved in the pandemic, that the cases are sporadic and even in case of a wider diffusion Finland has well enough human resources and sanitary devices to face it. The epidemiological situation is being monitored at the regional and national levels. Regional and local authorities decide on introducing or lifting measures according to the epidemiological situation. On 6th May 2020, the Government adopted a resolution implementing a plan for a hybrid strategy to manage the COVID-19 crisis. The plan aims to prevent the spread of the virus in society, to safeguard the carrying capacity of the healthcare system and to protect people, especially those in risk groups. The action plan for implementing the hybrid strategy, adopted in autumn 2020, guides the authorities in managing the COVID-19

[7] <https://stm.fi/koronavirusepidemian-tasot> (06/2021). The translations in the text are by the authors if not otherwise indicated.

epidemic. It sets out the criteria used to determine the phase of the epidemic. The action plan has been updated for January-May 2021.The action plan was supplemented with the new tiers of prevention measures on 26th January 2021. In its resolution on 6th May 2021, the Government stated that the use of nationwide tiers of COVID-19 prevention measures could be discontinued. On 16th March, the Finnish Government, jointly with the President of Finland Sauli Niinistö, declared a state of emergency due to COVID-19. 272 laboratory-confirmed cases caused by COVID-19 had been diagnosed in Finland by 16th March 2020 at 2 PM. The head of THL, Markku Tervahauta, told MTV3 that the actual number of COVID-19 cases might be 20-30 times higher than what had been confirmed by testing, due to the fact that testing was limited to risk groups, the severely ill, and healthcare workers. On 16th March, the Government also announced they had decided to take the following measures by issuing a decree on implementing the Emergency Powers Act (*valmiuslaki*). The measures were scheduled to be in place until 13th April, after approval by the Parliament of Finland, but were later extended to 13th May. The state of emergency was dismissed on 16th June 2020.

- All schools are closed, not including early education.
- Most government-run public facilities (theatres, libraries, museums etc.) are shut down.
- Critical personnel are exempted from the Working Hours Act and Annual Holidays Act, both in the private and public sector.
- At most 10 people can participate in a public meeting, and people over the age of 70 should avoid any human contact whenever possible.
- Outsiders are forbidden from entering healthcare facilities and hospitals, excluding relatives of critically ill people and children.
- The capacity of social and healthcare will be increased in the private and public sector, while less critical activity will be decreased.
- Preparations for the shutdown of borders will start, and citizens or permanent residents returning to Finland will be placed under a two weeks quarantine.

The political and governmental systems differ in Finland and in Italy in many ways. In particular, in Italy every new set (or loosening) of restrictions was given through a decree while in Finland such kind of decrees do not exist and the government has provided (and is still providing) suggestions and recommendations instead of Italian-style imperatives. In Finland, in the worst moment the state of emergency (*poikkeusolot*, 13th March 2020) has been declared and consequently the Emergency Powers Act came into force. Through it a series of restrictions and economic aids have become possible and have been implemented.

Since the beginning of the COVID-19 emergency both Italian and Finnish governments have released official communications on weekly basis (or even daily in some moments). Between the beginning of the health emergency and the present writing 122 official communications have been released by the Finnish government and 24 decrees by Giuseppe Conte.

3. Information pollution

The uniqueness of this unprecedented event significatively emerged in communicative responses as well. The COVID-19 pandemic has been and still is a problem that goes beyond the medical-epidemiological field. In fact, since its emergence, this issue has also contributed to, and has been accompanied by, a global spread of another type of pandemic phenomenon, but of an informative nature, equally dangerous and difficult to control. In mid-February, the

World Health Organization announced that the new coronavirus pandemic was accompanied by an *infodemic* of misinformation (2020).[8]

In an information society like the current one, the very high media coverage and the diffusion of news, statistical data and live scientific results, have produced an unprecedented social, anthropological and communicative emergency on a worldwide scale. This information mode, made possible and strengthened by algorithms underlying the new media and their own design, conceived to maximize the dissemination of information with a virtual *cascade effect*, no longer taking into account the complexity of the information as a parameter of reference. Moreover, all the subjects (citizens, stakeholders, companies, scientists and politicians) are using the same communication tools sharing the same platforms which are all owned by private companies (e.g. Facebook, Twitter and Google), often with polarized outcomes. As suggested in Crocchi (2020), the functioning of the current news market perpetuates the polarization of news, but the area of popular concern has recently focused on exposure to ideological difference in online media. Social networks and online communication have considerably transformed political communication and public discourse. Over the past two decades, changes in communication technologies have affected the very nature of how citizens relate to each other and to their governments, shaping a new public sphere (*ibidem*: 397).

At the time of COVID-19, understanding this *information pollution* is determining to more effectively manage the practical and concrete interventions that the emergency requires. Information pollution in its broader sense implies a large spectrum of phenomena which may affect communication at various levels, manifesting themselves as a unicum or simultaneously with others. The direct and indirect impacts of information pollution are difficult to quantify but there is undoubtedly much evidence of it in the narrative of the COVID-19 so far. For the aim of our discussion, we consider the conceptual framework for examining information disorder introduced by Wardle and Derakhshan (2017) where they identify three types of information disorders (Dis-information, Mis-information and Mal-information) distinguishing elements (agent, message and interpreter), phases (creation, production and distribution) and purposes among them. Much of the discourse on *fake news*, for instance, conflates three notions: mis-information, disinformation and mal-information.[9]

Today the complexity of communication is so manifold that it is unrealistic to attempt to describe it in an exhaustive way. Among all the possible perspectives, at least two seem significant for the aim of this paper. The first is the perspective of the sender. In this context, it seems legitimate to distinguish between institutional figures, communication professionals (e.g. journalists and opinion leaders) and a third macro-category "other" that does not fall within the previous two and include all non-official channels of information. The second perspective inevitably concerns the media used. Nevertheless, there are at least three possibilities: institutional media (the website of the Government, the Ministry of Health, etc.), media outlets (newspapers, television broadcasts, etc.), social media (in the broader meaning: Facebook, Twitter, Youtube, etc.). Social media represent great novelty compared with the previous global emergencies as they also play the crucial role of filters (e.g. filter bubbles) or echo chambers (Pariser 2011;

[8] <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19> (06/2021).

[9] They distinguish between the three disorders as follows: *Mis-information* is when false information is shared, but no harm is meant. *Dis-information* is when false information is knowingly shared to cause harm. *Mal-information* is when genuine information is shared to cause harm, often by moving information designed to stay private into the public sphere.

Bakshy 2015). These two perspectives are inevitably intertwined, although in different ways. For example, a speech by the Prime Minister can be disseminated both from the institutional website of the Government, and from the press, and from social media and social networks. On the contrary, a post on Facebook by a non-institutional or professional figure will hardly go beyond the confines of social networks. But that doesn't mean it will have a lower spread.

The development of e-Government and the availability of online information for citizens varies considerably across the EU Member States. According to a 2020 Eurostat survey, the percentage of regular web users who refer primarily to institutional communication channels in Italy, is significantly lower than that of other European countries, such as Finland, Germany or France.[10]

People increasingly use the internet to carry out their daily tasks. Over the past years, the use of e-Government has also gained popularity as it enables citizens to obtain information at any moment or carry out administrative tasks remotely. In 2020, 47% of people in the EU aged 16-74 had obtained information from the websites of public authorities during the last 12 months prior to the survey. This share increased substantially compared with 2008 (33%). Citizens of all ages use public authorities' websites to obtain information. The share of people that reported to have done this in the last 12 months was highest among the 25-34 years old (59%), followed closely by those 35-44 years old (56%). Although older citizens were less likely to use public authorities' websites to get information, more than a quarter (26%) of citizens aged 65-74 had used government websites for this purpose in the last 12 months. In 2020, the highest shares of people that used the websites of public authorities to obtain information in the last 12 months were recorded in Denmark (89%), Finland (85%), the Netherlands (81%) and Sweden (79%). In contrast, only 10% of citizens in Romania had obtained information from public authorities through their websites. This was also not common in Italy and Bulgaria (both 19%).[10]

Many surveys have been conducted after the spread of the pandemic and the global experience of lockdown and restrictions, and at the time of writing, some significant data regarding media consumption and trusted sources of news emerge. According to a survey from April 2020, Italian people considered TV newscasts the most reliable news source regarding the coronavirus (COVID-19). The Government followed in the ranking with 48% of individuals seeing it as a reliable news source. News shared by friends and family were perceived as more reliable (20%) than radio (17%).[11]

[10] <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210306-1> (06/2021).

[11] All data in the report comes from a GlobalWebIndex custom survey conducted between March 31st and 2nd April, which tried to gather at least 1.000 Internet users, aged between 16 and 64, for 17 countries including: Australia, Brazil, China, France, Germany, Italy, Japan, Philippines, Singapore, South Africa, Spain, UK and USA, India, Canada, New Zealand, Ireland. The answers were collected based on age, gender and education. In some countries, a small number of responses were found for groups 16-24 or 55-64 in these cases responses were associated with adjacent age groups. For this wave 2 the Italian champion is made up of 1.075 users.

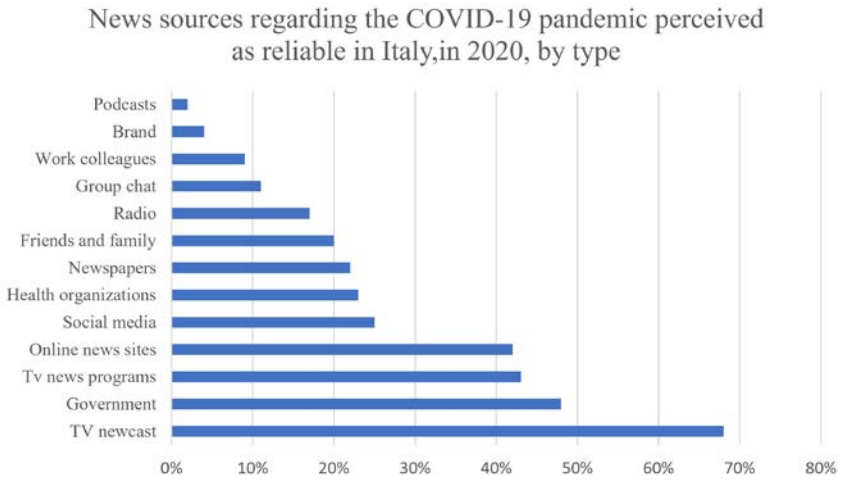


Table 1. News sources regarding the COVID-19 pandemic perceived as reliable in Italy, in 2020, by type (Statista 2020a)

As for the most trusted sources of coronavirus news in Finland in the same year, television news and newspapers were considered the most trusted sources for information about the coronavirus (COVID-19). For around 70% of Finns, TV news was the most trusted news source, followed by newspapers with a share of 60%. One third of the respondents had the highest confidence in radio news, while other news sources, especially online, were not considered as trustworthy.

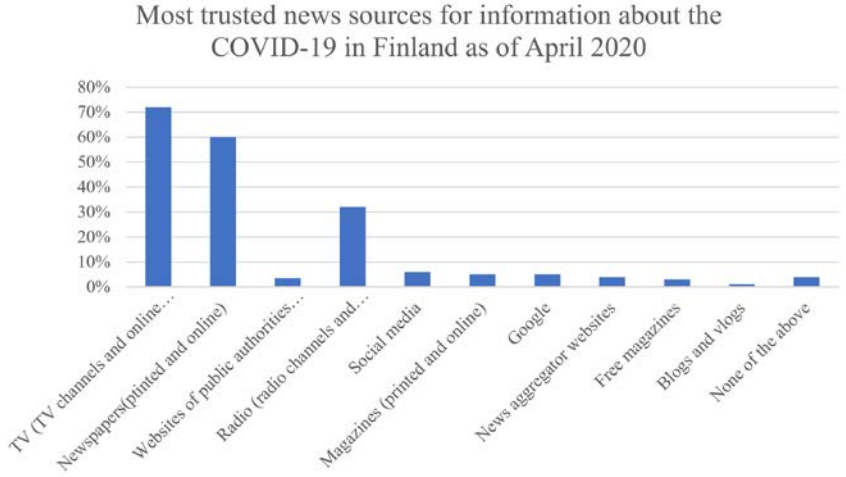


Table 2. Most trusted news sources for information about the COVID-19 in Finland as of April 2020 (Statista 2020b)

The most significant element that emerges from these surveys is that the audience (intended as recipients of the message) perceives social media as less reliable for receiving truthful information on the progress of the pandemic; in fact, within the broad spectrum of information disorders it is possible to trace the greatest responsibility of spreading fake news to social media (and its users). The pervasive use of social media and the spread of misinformation among them forced

some governments, such as the Italian one, to adopt hybrid communication strategies where, for instance, the official statements of the Prime Minister were announced and posted on his official Facebook profile (see section 6). For the aim of this study we analyse the institutional communication intended as communication realized in an organized manner by an institution or its representatives with the main aim of setting a good relationship between the institution and its potential public (Lever *et al*. 2002). It aims to establish quality relationships between the institution and the public with which it relates, to achieve social notoriety and adequate public image for the purposes and activities of the institution itself in terms of Public communication (La Porte 2021). Institutional communication, as mentioned before in the paragraph, is not immune to certain information pollution as well. For instance, at the beginning of March 2020, the Italian government was involved in a serious episode of *Mal-information*, defined as the genuine information that is shared to cause harm. This includes private or revealing information that is spread to harm a person or reputation (Wardle and Derakshan 2017). The draft of the government decree of the 8th March 2020 with the new strict measures to fight the coronavirus ended up in the main online newspapers – in Italy and in the world. According to the CNN online newspaper, that document was provided to them by the press office of the Lombardy Region who denied any involvement. This serious episode of news leaks caused a real crisis communication example within the broader emergency scenario caused by the pandemic.

4. Lexical choice in the discourse of the Italian and Finnish governments: extracts of two speeches

In what follows two communications of the Italian and Finnish governments and the day-after headings on two national newspapers are analysed. The speeches are given by the then-Prime Minister Giuseppe Conte (2018-2021) for Italy and by the Prime Minister Sanna Marin (2019-present) for Finland and are representatives of two crucial moments in the COVID-19 pandemic during 2020. The first one is the first available official communication on the spread of COVID-19, and the second one is representative of the second stage when the "cohabitation with the virus"-period started. In the discourses we paid attention in particular to the lexical choice of verbs, nominals, adjectives and adverbs and in how the discourse is overall construed taking into account the six factors presented in section 1.

On the basis of the audience to which it is addressed, the genre of these discourses can be considered as both informative and persuasive: informative about the ongoing health emergency and persuasive as they aim at convincing people to comply with the restrictions and recommendations provided by the prime ministers.

	Date	Approximate word count[12]
Italy	22.2.2020	863
	26.4.2020	3.479
Finland	27.2.2020[13]	2.093
	4.5.2020[14]	304

Table 3. Government communications and word count

[12] The greetings at the beginning as well as the questions of the journalists and the answers to them are excluded from word count and analysis.

[13] <https://valtioneuvosto.fi/-/hallituksen-tiedotustilaisuus-koronavirustilanteesta-ja-varautumisesta-suomessa-klo-13-30> (06/2021).

[14] <https://valtioneuvosto.fi/-/10616/hallitus-linjasi-suunnitelmasta-koronakriisin-hallinnan-hybridistrategiaksi-ja-rajoitusten-vaiheittaisesta-purkamisesta> (06/2021).

We then observe coronavirus-related headings on the day-after newspaper (one for each country) to check how many of them concern the official communication. We aim at comparing the language, namely, whether the lexical choice is similar or not, whether words with a greater impact are used or not, whether other kinds of strategies, such as the choice of verbal forms, are adopted. We focus on the news headings in particular as we consider them to represent on the one side the core of the article and on the other hand an invitation to further reading and thus having necessarily a persuasive impact on the readers (Santulli 2015).

Both Italian and Finnish governments have provided communications on the evolving situation on a regular basis since the end of February 2020. The communications are given by representatives of the governmental and/or sanitary authorities and during every communication more than one person speaks and presents the situation of her/his competence. In the present paper we restrict our analysis on the discourse of the prime ministers Giuseppe Conte and Sanna Marin.

4.1 Italy: Communication 1, 22.2.2020

The first official communication of the Italian government regarding coronavirus was released on 22nd February 2020 and it is followed by other two talks, respectively, from Roberto Speranza (Minister of Health) and Angelo Borrelli (head of the Civil Protection Department). The talk is divided into three parts.

In the first part, Conte stressed the importance of the efforts of all the parties involved, the use of the language is shaped around the concept of the meticulousness of the commitment carried out by everyone. It is noticeable how he wants to highlight the teamwork and the scrupulousness of every step taken by the Italian government by some lexical choices such as the verb *impegnarsi* 'to commit' and the corresponding nominalized form *impegno* 'commitment' and the repetition of the verb *lavorare* 'to work' and *tutelare* 'to protect'. In the second part Giuseppe Conte spends some minutes thanking many institutional figures stating very clearly that all of them are involved in the containment of the virus. The element that emerged mostly is the desire to make people perceive that the work was capillary and that it was carried out with the utmost caution. The president emphasizes the actions taken by Italy underlining the caution, the methodological rigor and the timeliness of certain choices (Italy was in fact the first country to block air traffic to and from China) through the use and the repetition along the talk of superlative adjectives as *massimo* 'best, utmost' (6 times). In the third part he accentuates the urgence of take imminent actions in the near future stressing the importance of moral rigour of the citizens by lexical choices such as: *autoresponsabilità* 'self-responsibility', *autocensura* 'self-censorship', *comportamenti* 'behaviors'(x 3) . The use of adverbs are pretty noticeable: *sicuramente* (x2) 'surely', *costituzionalmente* 'constitutionally', *ovviamente* (x2) 'obviously', *assolutamente* (x2) 'absolutely' and *sempre* (x4) 'always'. In this first intervention there is only one direct reference to the *Coronavirus,* instead the noun *emergency* is used twice. What stands out the most along the talk is the massive usage of the first plural form of every verb to convey this sense of unity that we mentioned before. In the bipartition of inclusive (me + you) and exclusive (me + s/he or them) "we" (Filimonova 2005) the Italian Prime Minister uses mostly the exclusive "we" which gives a stronger leadership flavour to the whole speech and we can see behind it a willingness to increase trust to the action of the government. However, exclusive and inclusive "we" can appear even in the same sentence as in 3) below in which we have an esortative use *non dimentichiamo* 'let's not forget' and an exlusive "we" where the subject refer to the cabinet. The examples in Table 4 are directly drawn from the speech.

1) Abbiamo lavorato sino ad ora, è tutta la giornata che siamo impegnati per una riunione […] abbiamo convocato […]	We have worked untile now, we have been busy the whole day for a meeting […] we have summoned […]
2) […] abbiamo adottato un decreto legge […]	[…] we have adopted a decree […]
3) […] non dimentichiamo che stiamo parlando […]	[…] let's not forget that we are talking […]

Table 4. Clusivity of "we" in Conte's communication 1

In relation to the six factors presented in Section 1, the predominant ones are n.1 and 4: 1) *Harm-reducing, honest and accurate information*, which is explicitly stated by Giuseppe Conte during the talk. Lexical choices also contribute to reassuring the audience since he avoids fear inducing words (e.g. *disease*, *pandemic*, *epidemic*); the choice of using positive and constructive words and the constant use of first person plural for verbs also convey the sense of community and unity; 4) *The provision of proof of tangible actions* also is fulfilled by the explanation of what is possible to do in case of emergency, in fact Conte and his ministers are presenting the ten rules to avoid the spread of the virus and, moreover, he wants to underline the efficiency of the Italian health system. He clarifies the reason why Italy presents a high number of infected, and what is going to be done through the new decree.

4.2 Press in Italy 24.2.2020

In the eight COVID-19 related main headings of the day-after newspaper *Corriere della Sera*, one of the most widespread newspapers in Italy, no one explicitly refers to the speech of Giuseppe Conte the day before. The headings refer to many socio-economic aspects caused by the COVID-19 emergency. In 1) and 3) the titles focus on the restrictions and closures in Northern Italy. In 2) there is a short title addressing the food issue in supermarkets. The title in 4) expresses with very few words the most urgent aspect of the emergency: the need for intensive care units. This simple and short title conveys all the dramaticity of the situation. In 7) the topic is one of the points stated by Giuseppe Conte the day before during his talk, Italy (at that time) has the highest number of cases and there is an objective necessity to explain why.

During the pandemic there has been a large use of war metaphors and military jargon both in political communications and among the media, in general. The metaphorical presentation of the COVID-19 emergency as a war is not a new invention. There are several examples and studies in scientific literature that indicate how war metaphors are employed in everyday life (e.g. in politics, business, sport, love and moreover, disease). As stated in a recent study conducted by Panzeri, Di Paola and Domaneschi:

> The use of war metaphors is a widespread strategy in public speech for framing and representing the challenges to be faced. US Presidents declared war against poverty (Johnson, 1964), crime (Johnson, 1965), drugs (Nixon, 1971), cancer (Nixon, 1971), inflation (Ford, 1977); but this metaphor is so ubiquitous that it can also involve apparently not belligerent enemies such as traffic jams, sunshine and even salad 9. A study examined all articles published on three magazines (TIME, Newsweek and the Canadian Maclean) over a period of twenty years (1981–2000) and found that war/battle metaphors occurred in 15% of the articles. (2021: 2)

This powerfully evocative association might be dangerous because it could affect the way people conceptualize the pandemic and react to it, leading citizens to endorse authoritarianism

and strict limitations to civil liberties. The idea that conceptual metaphors actually influence reasoning has been corroborated by Thibodeau and Boroditsky (2011), who showed that, when crime is metaphorically presented as a beast, readers become more enforcement-oriented than when crime is metaphorically framed as a virus. Recently, Steen, Reijnierse and Burgers replied that this metaphorical framing effect does not seem to occur and suggested that the question should be rephrased about the conditions under which metaphors do or do not influence reasoning (2013). Nevertheless, the relationship between metaphor and interpretation is very close, especially in political communication. What defines the metaphorical use is therefore the connection between two lexical semantic contents and the context in relation to which they activate the search for further information, in particular in relation to the speaker. As stated in Baldi (2021), in political discourse the function of metaphor can be decisive: since it draws on two domains that correlate abstract notions to our experience of concrete realities, it is an effective way of making an abstract ideology accessible as an affective one.

Although we have not found in our analysis this type of use in the speeches of Conte and Marin, in the Italian news headings we can trace at least these words belonging to the semantic field of war: *sorvegliato speciale* 'special surveillance', *coprifuoco* 'curfew', *campo minato* 'minefield', *prigionieri* 'prisoners', *check point* and *nemico* 'enemy'. Also, in 1) words such as *chiusure* 'closures', *blocchi* 'blocks', *stop* and *no* contributes to convey a deep sense of anxiety which therefore clashes with the parameter 1) *Harm-reducing, honest and accurate information*, which instead, emerges from the institutional strategy in the speech of Giuseppe Conte the day before. Beside these lexical choices which are part of a clear communicative strategy used to convey a sense of anguish and emergence, the tone of the titles is pretty neutral except for 5) *La forza di reagire* 'the strength to react' which is more dramatic.

1) **Virus al Nord, chiusure e blocchi. L'emergenza: oltre 150 casi in 5 regioni. A Crema la terza vittima. Il nostro paese diventa "sorvegliato speciale", raddoppio di contagi in un giorno.** Stop a scuole, musei, cinema e pub. No ai turisti in Duomo. E l'Austria ferma per ore un treno dall'Italia.	Viruses in the North, closures and blocks. The emergency: over 150 cases in 5 regions. The third victim is in Crema. Our country becomes "under special surveillance", doubling infections in one day. Stop at schools, museums, cinemas and pubs. No tourists in the Cathedral. And Austria stops a train from Italy for hours.
2) **Corsa alle scorte. Gli scaffali vuoti nei supermercati.**	Stock rush. Empty shelves in supermarkets.
3) **E a Venezia addio Carnevale.**	And in Venice goodbye Carnival.
4) **Pronti 3.500 posti letto.**	3.500 beds ready.
5) **La forza di reagire.** Milano che chiude e annuncia il coprifuoco è l'immagine rovesciata di sé stessa, la città che appartiene alla gente si ferma nell'anticamera della paura, smarrita, quasi rassegnata, accettando una prova che mette a rischio la tenuta di un sistema. Il coronavirus è un campo minato da attraversare con prudenza e con ogni precauzione, ma i divieti alzano barriere mai viste e sperimentate in tempo di pace.	The strength to react. Milan that closes and announces the curfew is the inverted image of itself, the city that belongs to the people stops in the antechamber of fear, lost, almost resigned, accepting a test that jeopardizes the stability of a system. The coronavirus is a minefield to be crossed with caution and with every precaution, but the bans raise barriers never seen and experienced in peacetime.

6) Cibo a domicilio e check point. Vita da isolati. La zona rossa del Lodigiano. Le forze che presidiano. Palette e divise. All'interno ci sono circa 47 mila persone alle quali è stato consigliato di rimanere in casa, ma che volendo possono uscire per le vie del proprio comune o raggiungere gli altri paesi della zona rossa. Tutti "prigionieri" della quarantena.	Food delivery and check point. Life in isolation. The red zone of the Lodi area. The presiding forces. Palettes and uniforms. Inside, there are about 47,000 people who have been advised to stay at home, but who, if they want, can go out through the streets of their own municipality or reach other towns in the red zone. All "prisoners" of the quarantine.
7) **Milano. Un caso al Policlinico e molti uffici: lavorate da casa.** Perché da noi così tanti malati?	Milan. A case at the Polyclinic and many offices: you work from home Why are there so many sick with us?
8)**Le ricette globali.** Abbiamo visto con piacere il ministro della Salute e il presidente della Regione Lombardia annunciare insieme i dolorosi provvedimenti diretti ad evitare l'estensione del contagio. E letto con soddisfazione che anche il presidente della regione Veneto ha firmato insieme con il ministro la relativa ordinanza (la legge del 1978 sul Servizio sanitario nazionale richiede solo la firma del ministro). Ma, per affrontare questo nemico non basta la cooperazione nazionale.	8) The global recipes. We were pleased to see the Minister of Health and the President of the Lombardy Region jointly announce the painful measures aimed at avoiding the spread of the infection. It is read with satisfaction that the president of the Veneto region has also signed the relative ordinance together with the minister (the 1978 law on the National Health Service requires only the signature of the minister). But national cooperation is not enough to face this enemy.

Table 5. COVID-19 related headings on CdS, 24.2.2020

4.3 Italy: Communication 2, 26.4.2020

The second communication under analysis was released on 26th April 2020 at the beginning of a second stage of the emergency due to coronavirus. Italy is now moving to the so-called *Fase due* 'Second stage'. This a long talk made by Giuseppe Conte only, crucially there is no intervention of other institutional representatives. The state of emergency is still valid even if the spread of the virus is more under control. The length of the talk is mainly due to the complexity of the new norms which will be valid from 4th May 2020 on. In the first part of the talk, he underlines the complexity of the situation praising the efforts made by the citizens to contain the virus. There is a pervasive use of the verbs: *monitorare* 'to monitor', *controllare* 'to control', *verificare* 'to verify', the explicit subject of all these verbs is the Italian Government. In the first part of this talk, Conte makes a sort of claim (intended as in marketing and commercial sense) saying: *Se ami l'Italia, mantieni le distanze* 'If you love Italy, keep the distances'. This is quite unusual in the context of oral institutional communication but it may be connected to the pervasive use of (personal) social media and the tendency to homologate the syntactic choices in all his communicative channels. As we are going to analyse later in the paper, the communication strategies adopted by Giuseppe Conte and his staff are different from the one of the Finnish Government and this specific choice may also affect his language. There are not significant recurrences in lexical choices except for some nouns such as *salute* 'health' (x6), *protocollo di sicurezza* 'security protocol' (x4) and *comitato tecnico-scientifico* 'technical-scientific committee' (x4). The pervasive use of many modal verbs with the form *dobbiamo* 'we must'

(x13) is what stands out the most, the Prime Minister uses *dobbiamo essere consapevoli* 'we must be aware', *dobbiamo rendercene conto* 'we must realize', *dobbiamo rimboccarci le maniche* 'we must roll up the sleeves', *dobbiamo assumerci* 'we must take on', *dobbiamo rispettare* 'we must respect', *dobbiamo approfittare* 'we must take advantage', *dobbiamo continuare* 'we must continue', *dobbiamo sbloccare* 'we must unblock', *dobbiamo essere orgogliosi* 'we must be proud'. All of these verbs are used with the first plural form with the clear goal to convey a sense of unity and common civil effort that must be committed by everyone. Hence, in this speech we can observe a clearly predominant use of inclusive "we". Whereas in the first speech the Prime Minister placed himself among the authorities, being the utmost authority speaking to the country, now he is a citizen among the others, "one of us". In 4) below we can clearly see this dislocation: *bisogna rispettare* 'it must be respected/we must respect' is an impersonal form that includes the speaker as well his audience, as well as the expression *teniamo conto* 'let's take into consideration' an esortative first person plural, and the scientists and experts are a third part, 'they'. In 5) the inclusive "we" is even more stressed by the pronoun *tutti* 'all'.

4) […] bisogna rispettare questa precauzione teniamo conto che gli scienziati e gli esperti ci dicono […]	[…] this precaution must be respected, let's take into consideration that the scientists and experts tell us […]
5) vogliamo tutti che il paese riparta […]	We all want the country to restart again […]
6) potremmo affidarci a risentimento […] potremmo prendercela con chiunque capiti a tiro […]	we could rely on resentment […] we could go after anyone who happens within range […]

Table 6. Clusivity of "we" in Conte's communication 2

As for the second talk of Giuseppe Conte, the main points are n.1, n.4 and, above all, n.3: 1) *Harm-reducing, honest and accurate information*; as for this point in this speech made by the Prime Minister Giuseppe Conte emerges, above all, the honesty with which he describes the efforts made by the Italian people during the first phase. 3) *Clear objectives and goals and how they are presented*; the main purpose of this talk, in fact, is the presentation of the second phase goals and new rules. The Prime Minister is very clear about the risks that this new phase may cause. He expresses his desire to grant to everyone a few days of preparation in view of the effective entry into force of the second phase's rule which are anticipated by a clear 4) *provision of proof of tangible actions.*

4.4 Press in Italy 28.4.2020

As for the communication given on 27.4.2020 and the related day-after newspaper there is only one mention of the Prime Minister's speech in title 1) The tone is neutral and it is just a statement on the new rules and prohibitions mentioned in the speech the day before. The sense of objectivity is given by the use of reported speech. Interestingly, in the subheading the quoted part is the last one after a brief list of the changes in the restrictions and it contains the exhortation to continue to be cautious. The fact of being quoted highlights this part and it is thus stressed as the most important information (Antelmi 2012 among others).

Four headings out of eight are related to economic matters such as financial help, funds, cuts and worries. The economic situation is the predominant theme over the whole narrative. The Chinese case is mentioned twice on the front page of the newspaper in two headings and the tone of the discussion in 2) and 8) is pretty polemic and controversial. In 2) there are clear

allegations against Chinese government alluding directly to a disinformation campaign perpetrated by Chinese authorities. In 8) the tone of the title and the discussion is both polemic and, for the first time here, personal. The author adds personal information and consideration to his piece; he also adds parenthetically a sort of request in the title: *per favore* 'please'. There are no explicit elements that allude to war related language.

1) **Si riapre con regole e divieti. Annuncio di Conte: i negozi riaprono il 18 maggio. Niente messe: l'ira dei vescovi.** Bar, ristoranti e parrucchieri: via libera il primo giugno. Esame di maturità in presenza. "Contagi ancora alti, serve cautela".	**We reopen with rules and prohibitions. Conte's announcement: shops will reopen on May 18th. No masses: the anger of the bishops.** Bars, restaurants and hairdressers: green light on 1st June. Graduation exams in presence. "Contaminations still high, caution is needed".
2) **Sulla Cina troppe ambiguità.** Scrive il New York Times, uno dei quotidiani più ostili a Trump dell'intero pianeta, che la Commissione europea avrebbe attenuato, su pressione dell'autorità di Pechino, un rapporto ufficiale sulle mistificazioni cinesi in tema di diffusione del Coronavirus. In questo rapporto si scriveva che "la Cina continua a condurre una campagna di disinformazione globale per sviare le accuse legate allo scoppio della pandemia".	**Too many ambiguities about China.** The New York Times, one of the most hostile newspapers to Trump on the entire planet, writes that the European Commission allegedly attenuated, under pressure from the Beijing authorities, an official report on Chinese mystifications regarding the spread of Coronavirus. In this report it was written that "China continues to conduct a global disinformation campaign to deflect allegations related to the outbreak of the pandemic".
3) **Il decreto in arrivo per le partite IVA. Il bonus salirà a 800 euro.** Bonus per gli autonomi che sale a 800 euro e di 600 per le babysitter. Un piano di circa 55 miliardi.	The incoming decree for VAT owners. The bonus will rise to 800 euros. Bonus for the self-employed which rises to € 800 and € 600 for babysitters. A plan of about 55 billion.
4) **Fondi diretti alle imprese.** Maurizio Casasco, presidente di Confapi: "Ora più risorse, ma i fondi vanno versati direttamente alle imprese".	Direct funds to businesses. Maurizio Casasco, president of Confapi: "Now more resources, but the funds must be paid directly to businesses".
5) **Tagli, infezioni e pochi tamponi: i medici passano alle denunce.** Mascherine inadatte, pochi tamponi, tagli: l'accusa di non aver tutelato chi lavora in ospedale. L'Italia è il Paese dove da anni la probabilità di prendersi un'infezione negli ospedali è la più alta: il 6 per cento. Oggi il personale sanitario – che conta 19.942 contagiati e 185 morti – attraverso i sindacati ha presentato un esposto ai Nas oltre che alle Procure di dieci regioni: contestano alle aziende ospedaliere di non aver tutelato medici e infermieri come dovuto.	Cuts, infections and a few swabs: doctors move on to complaints. Unsuitable masks, few swabs, cuts: the accusation of not having protected those who work in the hospital. Italy is the country where the probability of catching an infection in hospitals has been the highest for years: 6 percent. Today the health personnel - which counts 19,942 infected and 185 dead - through the trade unions presented a complaint to the Nas as well as to the prosecutors of ten regions: they criticize the hospitals for not having protected doctors and nurses as necessary.

6) **Mascherine, quando usarle. Fissato il prezzo: 0,50 euro.** Ci sarà un po' di allentamento sull'uso delle mascherine, anche se resteranno obbligatorie tutte le volte che non si riuscirà a rispettare un metro di distanza. E inoltre, sui mezzi pubblici, dal parrucchiere e nei negozi. Il prezzo? 50 centesimi.	Masks, when to use them. Fixed price: 0.50 euros. There will be some slack in the use of the masks, although they will remain mandatory whenever it is not possible to respect a distance of one meter. And also, on public transport, at the hairdresser and in shops. The price? 50 cents.
7) **Vittime in calo, la Lombardia torna al 7 marzo.** Ieri i morti in Italia sono stati 260. In Lombardia 56, il dato più basso dal 7 marzo. Sperando che questo sia il segnale che da giorni gli esperti si aspettavano di vedere. Ma nella regione che ha registrato ad oggi il maggior numero di decessi tornano però a crescere i contagi dopo sei giorni in discesa: l'incremento maggiore nell'area di Milano, più 463.Continua invece il calo di chi viene curato in terapia intensiva.	Victims in decline, Lombardy returns to 7 March. Yesterday there were 260 deaths in Italy, in Lombardy 56, the lowest figure since March 7, hoping that this is the signal that experts had expected to see for days. But in the region that has recorded the highest number of deaths to date, they return however, infections are growing after six days in decline: the greatest increase in Milan area, plus 463, while the decline in those who are treated in intensive care continues.
8) **La verità (per favore) su di noi.** Alla vigilia dell'8 aprile, quando è stato revocato il lockdown di Wuhan – un lockdown molto più rigido del nostro –, la Cina intera dichiarava 62 nuovi casi, la maggior parte dei quali importati. Il giorno precedente 32. Ieri, in Piemonte, la mia regione che non ho mai sentito così geograficamente lontana, i nuovi infetti confermati erano 394. Nella Lombardia limitrofa erano 920. Però apriamo. O meglio, iniziamo ad aprire, perché lo fanno anche gli altri, perché si avvicina l'estate e sotto sotto speriamo che il caldo ci dia una mano.	The truth (please) about us. On the eve of April 8, when the Wuhan lockdown – a much more rigid lockdown than ours – was lifted, the whole of China declared 62 new cases, most of them imported. The previous day 32. Yesterday, in Piedmont, my region that I have never felt so geographically distant, the new confirmed infected were 394. In neighboring Lombardy there were 920. But let's open. Deep down we hope that the heat will give us a hand.

Table 7. COVID-19 related headings on CdS 27.4.2020

4.5 Finland: Communication 1, 27.2.2020

The first official communication of the Finnish government regarding coronavirus was released on 27th February 2020. The first talk is by the Prime Minister Sanna Marin and it is followed by other two Ministers' talks. The communication is built around the core concept of being prepared to face the arrival of COVID-19 virus in Finland. Moreover, the ministers aim at reassuring citizens on the fact that in any case Finland is prepared to face the diffusion of COVID-19 in terms of medical competence and equipment, and for the moment there is no worry at all as the attested cases are just few and under control.

The core concepts of the discourse stand out from the lexical choice as can be inferred by the repeated use of the verb *varautua* 'to be prepared' and of the corresponding nominalized form *varautuminen* 'anticipation, the being prepared to'. As can be expected, the noun *koronavirus* 'coronavirus' is highly occurring even though it is not related to the word pandemic or emergency, which are never used. Also, nouns related to authorities such as *hallitus* 'government', *valtioneuvosto*

'cabinet' and *ministeriö* 'ministry' are frequent. However, the most used noun is *Suomi* 'Finland' with six occurrences. This, together with the use of first person plural for verbs, as in 7) creates a common ground, a sense of community based on the opposition of two entities: on the one side 'we, Finland' that is well-prepared to face something challenging and potentially dangerous, on the other side 'the other', which is not yet defined as an enemy but just as threat that is still far from the Finnish reality, together with 'the others', that is other European countries compared to which Finland has excellent medical and sanitary resources. The Finnish prime minister uses an inclusive "we" that refers to the country in its entirety. Hence, even if obviously who can raise the preparation of the country consist of the health and institutional authorities, it also includes the citizens who are called to respect the indications. In addition, we also find use of exclusive "we" that refers to the prime minister + other representatives of the cabinet and of the health authorities, as exemplified in 8). The examples in Table 8 are directly drawn from the speech.

7) […] ***pystymme kohottaa valmiuttamme nopeasti*** […]	[…] **we can** raise our preparation quickly […]
8) […] olemme halunneet asiassa noudattaa […]	[…] in this thing we wanted to obey […]

Table 8. Clusivity of "we" in Marin's communication 1

In addition, it is explicitly said that the official communications of the Finnish authorities are completely reliable. This again aims at creating a general sense of trust among citizens. Some adverbs in particular *tehokkaasti* 'efficiently', *tiivisti* 'tightly', *nopeasti* 'quickly', *jatkuvasti* 'continuously' can be semantically related to highlight the proper reaction of the government to something serious and unexpected (the spread of coronavirus) happening in a sudden and fast way but to which the government is reacting overall in an efficient way in addition to being in close collaboration with other entities and being continuously updated on the situation.

In relation to the six factors presented in Section 1. we can observe that:

1) *Harm-reducing, honest and accurate information*: this is explicitly stated by the Prime Minister and also emerges from the lexical choice in particular on two levels: on the one side by the avoidance of fear-creating words such as pandemic, epidemic, on the other side by the use of trust-creating words and use of first person plural for verbs.

2) *Simple and quick message*: this is clearly satisfied as both the speech itself and the sentences in it are quick and complex sentences (e.g. subordinate clauses, participial or infinitive phrases) are avoided. This is also in line with the recommendations of clear language to be used by the government.[15]

3) *Clear objectives and goals and how they are presented*, and

4) *The provision of proofs of tangible actions* also are fulfilled by the explanation of both what is possible to do in case of worsening of the general health situation and of what is going to be done right now (e.g. the sting of a COVID-19 coordination group formed by ministers and health authorities).

4.6 Press in Finland 28.2.2020

In the seven COVID-19 related main headings of the day-after newspaper *Helsingin Sanomat*, the national and most widespread newspaper in Finland, only one title refers to the speech of the

[15] <https://valtioneuvosto.fi/en/marin/government-programme/government-communications-strategy/values-in-government-communications> (06/2021).

prime minister Sanna Marin (n. 4 below). The headings report through the mean of indirect speech what is considered the core idea of the day before speech. Indirect or reported speech is generally used not only to vehiculate neutrality but also to give a scent of authority to the information, as discussed in Antelmi (2012) among others. In this title both the verb *varautua* 'to be prepared' and the corresponding nominalized form *varautuminen* 'anticipation, the (fact of) being prepared to' are used, which are exactly the same words used by the Prime Minister in her discourse. This also shows coherence between the institutional communication given on 27.2.2020 and the day-after news on the newspaper and thus fulfils factor 5) *Coherency among institutional communications and media (press)*. The other titles refer to chronicle (n. 1, 5, 7) or are comments or deepenings on the topic (n. 2, 3, 6). All the headings make use of present tense.

1) **Käsidesi käy kaupaksi juuri nyt, apteekista kerrotaan: "Pitää vain toivoa, että ihmiset käyttävät sitä myös".** Yliopiston Apteekin arjessa koronavirus näkyy vain vähän: suusuojainhylly on tyhjillään, ja osa asiakkaista haluaa jutella aiheesta. Ihmiset eivät vaikuta erityisen huolestuneilta, kertoo apteekinhoitaja Elina Lammi.	Hand sanitizer sells like hot cakes right now. From the pharmacy is said: "We have just to hope that people will also use it". In the ordinary day of the University Pharmacy coronavirus is barely seen: the mask shelter is empty and some of the customers want to talk about it. People don't seem particularly worried, says Elina Lammi.
2) **Voisiko koronavirus tarttua vierustoverilta 10 minuutin metromatkalla? Professori kertoo, miten lähikontakti määritellään.** Lähikontakti on THL:n mukaan henkilö, joka on ollut koronavirustartunnan saaneen kanssa tekemisissä kasvotusten tai samassa huoneessa yli 15 minuuttia. Määritelmä perustuu todennäköisyyksiin, sanoo epidemiologian professori Pekka Nuorti.	Can coronavirus be transmitted from the person sitting nearby in a 10 minutes subway trip? A professor tells how close contact is defined. According to THL, a close contact is a person who has been in contact with the coronavirus positive individual in the same space or room for more than 15 minutes. The definition is based on probabilities, tells the epidemiologist professor Pekka Nuorti.
3) **Hyvä käsihygienia torjuu koronavirusta, paniikki ei.**	A good hand hygiene defends from coronavirus, panic doesn't.
4) **Pääministeri: Koronavirukseen tärkeää varautua huolellisesti, mutta samalla välttää ylimitoitettuja toimia.** Sanna Marin antoi eduskunnalle pääministerin ilmoituksen Suomen varautumisesta koronavirukseen.	The prime minister: It is important to be prepared carefully for coronavirus but at the same time to avoid overscale actions. Sanna Marin gave the prime minister's announcement to the Finnish parliament on the anticipation of Finland to coronavirus.
5) **Ylilääkäri: Helsingin koronaviruspotilas voi hyvin, THL selvitti naisen lähikontaktit matkareitiltä.** Tartunnan saaneeseen suomalaisnaiseen matkareitillä lähikontaktissa olleet henkilöt ovat eri maiden viranomaisten tiedossa. Viranomaiset ovat heihin yhteydessä. THL ei erittele tarkemmin, mistä maista henkilöt ovat kotoisin.	The medical superintendent: The coronavirus patient in Helsinki is good, THL examined the contacts of the woman on the trip. The persons that the infected Finnish woman met on her trip are known by the authorities of the different countries. The authorities are in contact with them. THL does not give details on the countries of origin of these persons.

6) **Voiko koronavirus tyhjentää kauppojen hyllyjä Suomessa? Tietyistä älypuhelimista voi tulla pulaa jo lähiviikkoina, arvioi S-ryhmä.** Elintarvikkeisiin ja käyttötavaroiden tuontiin virus ei näytä vaikuttavan tänä keväänä juuri lainkaan, suurostajat arvioivat.	Can coronavirus empty shops' shelters in Finland? Some smartphones can be missing in next weeks, estimates the S-group. The virus does not seem to have an impact on food availability and import this spring.
7) **Ensimmäistä koronavirusrokotetta aletaan testata huhtikuussa.** Rokotteet valmistuvat aikaisintaan vuoden päästä. Lääkettä tautiin pitää odottaa kauemmin. Nykyisistä lääkkeistä on jo saatu lupaavia tuloksia.	The first coronavirus vaccine will be tested in April. The vaccines will be ready not earlier than one year. A medicine for the disease could take even longer. At present there already are promising medicines.

Table 9. COVID-19 related headings on HS, 28.2.2020

4.7 Finland: Communication 2, 4.5.2020

The second communication under analysis was released on 4th May 2020, when the beginning of a second stage of the health emergency due to coronavirus is identifiable also in Finland. The country is now moving towards the adoption of the so-called hybrid strategy which is based on four key actions: *testaa-jäljitä-eristä-hoida* 'test-track-isolate-cure'.

The spread of coronavirus is better under control and even if the state of emergency is still valid some of the restrictions can be now dismissed. This is evident in particular in the use of nominalized forms such as: *jatkaminen* 'continuation', *purkaminen* 'dismission'. The dismissal of the more restrictive indications should be done in a restrained and gradual way in order to control the diffusion of the virus and this is emphasized by the very frequent use of manner adverbials such as *hallitusti* 'restrainedly', *asteittain* 'gradually', *porrastetusti* 'step by step', and by nominalized forms such as *hillitseminen* 'control'. Interestingly, different from the first communication, in this one no use of first-person plural is made. However, the passive is widely used. This form can be used both with as an impersonal form (in Finnish in fact the agent cannot be expressed in this construction), as in 9), or for first person plural especially in colloquial Finnish, as in 10).

9) Kun rajoitustoimia [...] poistetaan, epidemian kehittymistä ja terveydenhuollon kuormittumista seurataan ja arvioidaan tarkasti.	As restrictions are dismissed […] the development of the epidemic as well as the overload of the health system are followed and evaluated carefully.
10) [...] voidaan asettaa rajoituksia [...] ulkotiloja voidaan avata hallitusti [...]	[…] it is possible/we can set restrictions […] outdoor spaces can be opened restrainedly […]

Table 10. Clusivity of "we" in Marin's communication 2

As for the factors presented in Section 1. we can observe that:

1) *Harm-reducing, honest and accurate information*: information is honest and accurate and it is in fact reported on the consultations of the cabinet and on what are the future indications of the government (a detailed list is given).

3) *Clear objectives and goals and how they are presented*: Goals are presented clearly, as evident from explicit statements such as *Hallituksen tavoitteena on…* 'The goal of the government is …'.

4) *Provision of proof of tangible actions* is visible from the statements about the intention to constantly control and evaluate the development of the epidemic and the related possible (over)load of health care. Moreover, it is also stated that *epidemian hillitsemisessa on onnistuttu Suomessa toistaiseksi hyvin* 'the control of the epidemic has been for the moment successful in Finland' a piece of information that instils self-confidence (and confidence towards the *modus operandi* of the government) in the population.

At the time of this speech, it is already clear that it's not 'just' a health emergency but it is a serious worldwide epidemic emergency that has now attained Finland, too. The reality is not diminished nor trivialized (accomplishing again to factor 1) and coherently the word *epidemia* 'epidemic' is used.

The general sense of the discourse is 'we have done good, now we can *gradually* dismiss some restrictions. Again, the message is quick and simple (factor 2), the core idea of 'gradually re-opening' is repeated and, as in communication 1, language is clear and complex structures are avoided.

4.8 Finland: newspaper 5.5.2020

As for the communication given on 4.5.2020, we find seven headings related to coronavirus in the day-after newspaper *Helsingin Sanomat*. Among them three titles are directly related to the communication given by the Government the day before (n. 1, 2, 6). Only title 1) contains reported speech. This gives a tone of authority to the reported information, as was the case for the reported speech in Table 9. The others are comments (n. 4) or a deepening of the current health emergency (n. 3, 5, 7). Overall, all the news related to the government's communication are coherent with it (factor 5) and no emphatic tone is used to exaggerate the information provided by the Prime Minister. Only in title 2 the sentence *hallitus ripustaa nyt Suomen neljän sanan varaan* 'the government hang up Finland on four words', which hints at the hybrid strategy put in place by the government and based on test-track-isolate-cure, can be interpreted as slightly ironic and sceptic. Finally, we interpret the choice of neutral and non-dramatic lexical items as being in line with a general approach of Finnish media based on reliability, transparency and clear communication that is coherent with the official channels.

1) **Suomi avautuu merkittävästi kesäkuusta lähtien – Pääministeri Marin varoittaa, että tilanteen muuttuessa rajoituksia pohditaan uudestaan.** Hallituksen johtoviisikko kertoi päätöksistä rajoitustoimien purkamiseksi tiedotustilaisuudessa maanantai-iltana.	Finland will open significantly from June on – the Prime Minister Marin warns that in case the situation changes restrictions will be re-thought. The five guiding the government told about the dismissal of the restrictions in the communication on Monday evening.
2) **Kesäkuussa alkaa koronakriisin suuri käänne, ja hallitus ripustaa nyt Suomen neljän sanan varaan.** Suomen päätavoite ei ole hävittää koronavirusta vaan hallitus rakentaa omaa suhteellisen avointa reittiään, kirjoittaa HS:n talouden ja politiikan toimituksen esimies Jussi Pullinen.	In June the big twist of coronacrisis will start and the government hang up Finland on four words. The main objective of Finland is not to cancel coronavirus but to build a relatively open way out, writes Jussi Pullinen.

3) **Koronakriisi teki epidemiologeista ja virologeista planeetan tunnetuimpia ihmisiä – HS esittelee viisi julkkistutkijaa, joiden päätökset voivat sulkea valtioita.** Tutkijoiden suositukset sulkevat nyt valtakuntia. Koronakriisi on repäissyt epidemiologit ja virustutkijat tutkijankammioistaan valokeilaan.	Coronacrisis puts epidemiologists and virologists among the most famous persons in the world. HS presents five famous researchers whose decisions can close states. The recommendations of researchers are now closing states. Coronacrisis has ripped epidemiologists and virologists from their offices in the beam of light.
4) **Yksittäisten päivämäärien sijaan on tärkeämpää kuvata rajoitustoimien muutosten logiikka, koska se luo uskoa tulevaan.** Jos tartuntatautilaki ei riitä epidemian tehokkaaseen torjuntaan, voisi olla parempi korjata tartuntatautilakia kuin turvautua uudelleen valmiuslain pykäliin.	Instead of single dates it is more important to describe the logic of the restrictions because it gives trust in the future. If the law on contagious disease is not sufficient it could be better to modify that law than to use the Emergency Powers Act.
5) **Kirjastojen ja museoiden avaamislupaus oli helsinkiläisten toiveissa jo etukäteen:"Askel kerrallaan kohti normaalia".** HS:n maanantaina haastattelemat kansalaiset antoivat tukensa esimerkiksi kirjastojen ja museoiden hallitulle avaamiselle.	The re-opening of libraries and museums was already on the wish list of people in Helsinki: "One step at time towards normality". The readers of HS that were interviewed on Monday supported the controlled re-opening of libraries and museums.
6) **Lue kattava lista siitä, miten rajoituksia nyt puretaan: ravintolat avautuvat asteittain, yli 70-vuotiaiden eristyssuositusta jatketaan.** Suomi avaa yhteiskuntaansa asteittain kesäkuun alusta alkaen.	Read a complete list of how restrictions are dismissed: restaurants reopen gradually, the isolation recommendation for over 70-years old is maintained. Finland re-opens its society gradually starting from June.
7) **Kaikki janoavat nyt tietoa siitä, miten epidemia etenee – THL:n mallinnukset ovat päätöksenteon ytimessä, mutta mihin ne perustuvat?** Epidemian mallintaminen ei ole matemaattisesti mahdoton tehtävä. Uuden viruksen kohdalla joudutaan kuitenkin tekemään paljon oletuksia. Se lisää ennustamisen epävarmuutta.	Everybody is eager to know how the epidemic will go on: the simulations of THL are the key for decisions but on what they are based on? The simulation of the epidemic is not mathematically impossible. However, many suppositions need to be put forth. This increases the uncertainty of the prediction.

Table 11. COVID-19 related headings on HS, 5.5.2020

5. Discussion

5.1 Communication and press 1

The Italian Prime Minister uses the noun *emergenza* 'emergency' whereas the Finnish Prime Minister never uses it and generally the communicative choice is centred on a more

reassuring tone. Nonetheless, the noun 'coronavirus' is used in both discourses marking the clearness and aiming at truthfulness and reliability. One striking difference is in the length of the communications: short and quick (but not less complete) as for the Finnish prime minister and longer and more articulated as for the Italian prime minister.

Comparing the day-after newspapers of the first communications, one Finnish newspaper's title refers to Marin's communication and in the Italian newspaper there is no explicit reference to Conte's communication even if one title concerns one of the main points faced by the Prime Minister the day before the official announcement. What overall stands out from the comparison of the Italian and Finnish newspapers' titles is the strikingly different tone that emerge in particular in the lexical choice: in the Italian ones there is recurring use of military or war terms (4 out of 8 titles) whereas in the Finnish ones the tone is much more neutral and less alarmist. Finally, only in two Finnish titles, one of which refers to the day-before speech, we observe reported speech whereas this mean of neutrality and objectivity (Antelmi 2012) is not used in the Italian titles, which on the contrary have in general a more dramatic tone.

5.2 Communication and press 2

The two communications are very different in length. This is mainly because of a different communicative strategy of the two governments. In fact, the Finnish prime minister basically resumes what will be explained in more detail by the ministers and health authorities afterwards, whereas the Italian prime minister is the one who carries out widely and in depth the whole communication of the government. Moreover, he makes use of repetitions and has a rhetorical tone speaking as he was alone the guide of the country in contrast with the first communication in which the facts, decisions and guidelines were presented as the result of a choral decision.

As for the day-after newspapers, in the Italian headings only one out of eight refers to Conte's speech, in the Finnish newspaper three headings out of seven. Interestingly, no war or military terms are now used. The attention is focused on the health emergency at a worldwide level and the lexical choice is more controlled and neutral. Only one title has a polemic flavour (title n.8) being a critical comment on the alleged lack of complete information on the pandemic.

The Finnish titles have a neutral tone as for the first set of titles. One title differs in this sense (n. 2), in which the tone is somehow sceptical of the hybrid strategy adopted by the government. In the Italian headings three out of eight use reported speech as a means of further authority and neutrality whereas in the Finnish ones only in one heading this strategy is adopted.

As a final note on the headings of the Italian and Finnish newspapers, we did not observe use of foreignism (excluded 'lockdown' in Italian) even if this is a commonly used strategy in newspaper headings.

6. Conclusions

In this paper we reported on the COVID-19 situation in Italy and in Finland with specific attention to institutional risk communication. In particular, after reviewing the respective national situations so far, we analysed four official communications of the Italian and Finnish governments, and specifically the speeches of the two prime ministers, in two important moments of the health emergency started in 2020: the very beginning and the end of the first stage when there is a gradual loosening of the previous restrictions. We compared the Italian and Finnish discourses and the day-after headings on two major newspapers from a communicative perspective grounding our observations on six parameters that we identified as crucial for

efficient institutional crisis communication. The analysis indicates that both prime ministers seem to have had a reliable and efficient institutional communication, even if communicative strategies differ for some aspects: the Italian prime minister Giuseppe Conte had much longer, articulated and rhetoric talks compared to the Finnish prime minister Sanna Marin. As for the day-after newspapers the Finnish headings maintain a much more neutral and factual tone whereas in the Italian ones sometimes use war and military jargon.

The element that emerges most significantly from the comparison of the communication strategy used by the two premieres is the different use of social media which helps shape the general communication strategy. Over the last decade, political communication has developed heavily on social media to implement strategies to increase consensus. The use of social networks by international and local politicians is crucial for both practical and strategic reasons. Social networks allow them to have direct contact with the electorate, being able to address different groups of voters simultaneously and with a number of diversified messages (according to the platforms used) is certainly one of the main advantages of political social media marketing. During the greatest spread of the COVID-19 pandemic in 2020, the communication plan has assumed an even more crucial value. World leaders had to manage an unprecedented situation that generated a sense of bewilderment among the citizens. During international crises, trust in government is expected to increase irrespective of the wisdom of the policies it pursues. This phenomenon has been called a *rally-round-the-flag* effect (Muller 1970). It explains the natural tendency of citizens to increase their trust in the political leader during times of crisis. This tendency was noted in moments of high tension such as the Cuban missile crisis or the attacks of 11th September 2001. Despite the different nature of this pandemic, the emergence of this tendency is – firstly – still noticeable and – secondly – amplified by a massive use of social media by political leaders. As for Italy, Giuseppe Conte has definitely focused on a communicative style emphatically centred on his figure: his press conferences were held at evening news time and they took place mainly on social media where he maintained his paternal and warm tone of voice. Some of the statement he pronounced sounded like commercial claims e.g. *Rimaniamo distanti oggi per riabbracciarci con più calore domani*[16] and the above-mentioned *Se ami l'Italia, mantieni le distanze.*[17] He used these phrases in the press conferences at first, and then they widely echoed from one social platform to another. Contrastingly, the choice to rely exclusively on institutional channels and leave to her Facebook profile a different role in the management of the crisis is perpetrated by Sanna Marin, despite the spread of the health emergency.

References

Albu, Oana B., and Mikkel Flyverbom. 2019. "Organizational Transparency: Conceptualizations, Conditions, and Consequences." *Business and Society*: 58 (2): 268-297.

An, Seonkyoung, and Karla Gower. 2009. "How do The News Media Frame Crises? A Content Analysis of Crisis News Coverage." *Public Relations Review* 35: 107-112.

Antelmi, Donella. 2012. *Comunicazione e analisi del discorso*. Novara: De Agostini.

Baldi, Benedetta. 2021. *Il Linguaggio del potere*. Roma: Carocci Editore.

Bakshy, Eytan, Solomon Messing, and Lada Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348 (6239): 1130-1132.

[16] Let's stay distant today to embrace each other more warmly tomorrow.

[17] If you love Italy, keep the distances.

Casalegno, Cecilia G., and Chiara Civera. 2016. *Impresa e CSR: La "non comunicazione" di successo. Regole per una gestione responsabile delle relazioni*. Milano: Franco Angeli.

Casalegno, Cecilia, Chiara Civera, and Damiano Cortese. 2020. "COVID-19 in Italy and Issues in The Communication of Politics: Bridging The Knowledge-behavior Gap." *Knowledge Management Research and Practice.* <https://www.tandfonline.com/doi/full/10.1080/14778238.2020.1860664> (06/2021).

Crocchi, Serena C. 2020. "Communication Strategies in The Infodemic Era." *Quaderni Di Linguistica e Studi Orientali*, 6: 383-404.

Dada, Sara, Henry C. Ashworth, Marlene J. Bewa, and Roopa Dhatt. 2021. "Words Matter: Political and Gender Analysis of Speeches Made by Heads of Government During The COVID-19 Pandemic." *BMJ Global Health* 6 (1): 1-12.

De Saussurre, Louis, and Peter J. Schulz. (eds.) 2005. *Manipulation and Ideologies in The Twentieth Century Discourse, Language, Mind*. John Benjamins Publishing Company.

Entman, Robert M. 1993. "Framing: Toward Clarification of a Fractured Paradigm." *Journal of communication* 43: 51-58.

Fischhoff , Baruch, and Julie S. Downs. 1997. "Communicating Foodborne Disease Risk." *Emerging Infectious Diseases* 3 (4): 489-495. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2640085/> (06/2021).

Jabeen, Sumera. 2016. "Do we Really Care about Unintended Outcomes? An Analysis of Evaluation Theory and Practice." *Elsevier Enhanced Reader. Evaluation and Program Planning* 55: 144-154.

Lakoff, George, and Mark Johnson. 2008. *Metaphors We Live by*. Chicago, IL: University of Chicago Press.

La Porte, José M. 2021. "Comunicazione istituzionale." ed. by Franco Lever, Pier C. Rivoltella, and Adriano Zanacchi. *La comunicazione. Dizionario di scienze e tecniche*. <www.lacomunicazione.it> (06/2021).

Leonard, Herman B., and Arnold Howitt. 2007. "Against Desperate Peril: High Performance in Emergency Preparation and Response." *Communicable Crises: Prevention, Response, and Recovery in the Global Arena*, ed. by Deborah E. Gibbons. Information Age Publishing.

Lever, Franco, Cesare Rivoltella, and Adriano Zanacchi. 2002. *La comunicazione: il dizionario di scienze e tecniche*. Cascine Vica: Elledici.

Menon, K.U., and Kelvin T. Goh. 2005. "Transparency and Trust: Risk Communications and The Singapore Experience in Managing SARS." *Journal of Communication Management* 9 (4): 375-383.

Mueller, John E. 1970. "Presidential Popularity from Truman to Johnson." *American Political Science Review* 64 (1):18-34.

Panzeri, Francesca, Simona Di Paola, and Filippo Domaneschi. 2021. "Does The COVID-19 War Metaphor Influence Reasoning?" PLoS ONE 16 (4): e0250651. doi: 10.1371/journal.pone.0250651.

Parisier, Eli. 2011. *The Filter Bubble: What The Internet Is Hiding from You*. New York, NY: Penguin Group.

Poirier, William, Catherine Ouellet, Marc-Antoine Rancourt, Justine Béchard, and Yannick Dufresne. 2020. "(Un)Covering the COVID-19 Pandemic: Framing Analysis of the Crisis in Canada." *Canadian Journal of Political Science. Revue Canadienne de Science Politique*: 1-7.

Reynolds, Barbara, and Matthew W. Seeger. 2005. "Crisis and Emergency Risk Communication as an Integrative Model." *Journal of Health Communication* 10: 43-55.

Santulli, Francesca. 2015. "La riflessione metalinguistica sulla stampa italiana: oltre l'epicedio?" *Circula* 2: 55-75.

Statista. 2020a. "News Sources Regarding the Coronavirus (COVID-19) Pandemic Perceived as Reliable in Italy in 2020, by Type." <https://www.statista.com/statistics/1113715/reliable-news-sources-during-the-coronavirus-italy/> (06/2021).

Statista. 2020b. "Most Trusted News Sources for Information about The Coronavirus in Finland as of April 2020." <https://www.statista.com/statistics/1114881/trusted-sources-of-coronavirus-news-in-finland/> (06/2021).

Steen, Gerard, Gudrun Reijnierse, and Christian Burgers. 2014. "When Do Natural Language Metaphors Influence Reasoning? A Follow-Up Study to Thibodeau and Boroditsky." PLoS ONE 9 (12): e113536.

Ulmer, Robert R., and Timothy L. Sellnow. 2000. "Consistent Questions of Ambiguity in Organizational Crisis Communication: Jack in The Box as a Case Study." *Journal of Business Ethic*s 25 (2): 143-155.

Thibodeau, Paul, and Lera Boroditsky. 2011. "Metaphors We Think With: The Role of Metaphor in Reasoning." PLoS One 6: e16782

Tiirinki, Hanna, Liina-Kaisa Tynkkynen, Markus Sovala, Salla Atkins, Meri Koivusalo, Pauli Rautiainen, Vesa Jormanainen, and Ilmo Keskimäki. 2020. "COVID-19 Pandemic in Finland - Preliminary Analysis on Health System Response and Economic Consequences." *Health Policy and Technology* 9 (4): 649-662.

Turcotte, Anne M., Idriss Ali Gali-Gali, and Valéry Ridde. 2021. "The Unintended Consequences of COVID-19 Mitigation Measures Matter: Practical Guidance for Investigating Them." *BMC Med Res Methodol* 21 (1): 28.

Wardle, Claire, and Hossain Derakhshan. 2017. "Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making." Brussels: Council of Europe.

Wicke, Philip, and Marianna Bolognesi. 2020. "Framing COVID-19: How We Conceptualize and Discuss the Pandemic on Twitter." In *Computation and Language (cs.CL), Social and Information Networks*, <https://arxiv.org/abs/2004.06986> (06/2021).

Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

Citation: (2021) Contributors. *Qulso* 7: pp. 345-352. doi: http://dx.doi.org/10.13128/QUSO-2421-7220-12059

Contributors

Linda Badan is Associate Professor of Linguistics at Ghent University. She obtained her PhD at University of Padova and research fellowships as the Chiang Ching-kuo Foundation for International Scholarly Exchange (Taiwan) and the International Fellowships for Experienced Researchers Fernand Braudel-IFER (École Normale Supérieure, Paris). She has been post-doc fellow at City University of Hong Kong, Leiden University, École des Hautes Études in Science Sociales (Paris). Her main research interests include theoretical linguistics at the interface between syntax and prosody and syntax and pragmatics, comparative syntax between Romance and Sinitic languages and the acquisition of Italian as second language.

Laura Bafile is Associate Professor at the University of Ferrara. Her research concerns theoretical linguistics, phonology and the grammar of Italian dialects.

Benedetta Baldi is Full Professor of Communication and Discourse Analysis at the University of Florence. She is President of the Degree Course in Human Sciences for Communication and of the Second Cycle Degree Course in Theories of Communication. Her research includes different fields of linguistics and pragmatics, specifically language teaching and communication theory. As for the second subject, she is concerned with communication processes, political language, the relationship between media and cognitive and pragmatic capabilities and social/cultural identity. Her latest books are *#Opinione immediate* (Pacini, 2018), addressing communication and social media, and *Il linguaggio del potere* (Carocci, 2021), addressing the persuasive and manipulative nature of the political discourse.

Sergio Conti is adjunct professor of Chinese language and linguistics at the University of Naples "L'Orientale", the University of Siena and Roma Tre University. His research interests lie in the field of Chinese as a foreign language acquisition and teaching, particularly vocabulary, formulaic language, and the development of pragmatic competence.

Serena Camilla Crocchi is a PhD Student in Linguistics at the University of Florence under the supervision of Professor Benedetta Baldi. Her main areas of interest are in the field of communication theory, discourse analysis and semiotics.

Deborah Cappelli is a PhD student in Linguistics at the University of Florence. Her main research interest focuses on L2 acquisition, interlanguage, language transfer and educational linguistic of similar languages, in particular in Spanish and Italian language. The PhD thesis deals with the morpho-syntactic aspects in the language acquisition of similar languages under the supervision of Professor Benedetta Baldi.

Lena Dal Pozzo holds a PhD in Cognitive Sciences from the University of Siena. At present she works as Finnish language instructor and lecturer of Finno-ugric languages at the University of Florence, and holds the national scientific qualification for associate professor. Her research interests include L2 acquisition, bilingual acquisition, L2 theoretical applications to language teaching, Finnish syntax.

Elisa Di Domenico is Associate Professor of Linguistics. PhD in Linguistics, she has held a research appointment at the National Research Council in Rome, a post-doc position at the University of Florence and a teaching position at the University of Siena. Her research focuses on linguistic theory, the syntax-discourse interface and multilingualism.

Ludovico Franco is Assistant Professor in Linguistics at the University of Florence. His main research interests are morphosyntax and neurolinguistics. He has published a monograph and several papers in international journals, among which *NLLT*, *The Linguistic Review*, *Glossa*, *Lingua*, *Probus*, *Transactions of the Philological Society*, *Language Sciences*, and *Canadian Journal of Linguistics*.

Paolo Lorusso is Research Associate at the University of Florence and a member of the NetS IUSS lab for neurolinguistics, computational linguistics, and theoretical syntax. He obtained a PhD in Linguistics at the University of Florence and a PhD in Cognitive Science at Autonomous and University of Barcelona. His research areas are psycholinguistics, morphosyntax, syntax-semantics interface and comparative syntax.

M. Rita Manzini is Professor of Linguistics at the University of Florence. She is the author of several books, including *Locality* (MIT Press, 1991) and (in collaboration with Leonardo Savoia) *I dialetti italiani e romanci* (ed. Dell'Orso, 3 vols, 2005), *A Unification of Morphology and Syntax* (Routledge, 2007), *Grammatical Categories* (Cambridge UP, 2011), *The Morphosyntax of Albanian and Aromanian Varieties* (De Gruyter Mouton, 2018). She is also the author of several articles for volumes and journals, including *Linguistic Inquiry*, *Natural Language and Linguistic Theory*, *The Linguistic Review*, *Lin-gua*, *Probus*, *The Journal of Linguistics*, *Transactions of the Philological Society*. Her research focus is generative syntax and Morphosyntax.

Alberto Peruzzi is Professor of Theory of Knowledge at the University of Florence. His core research concerns a category-theoretic framework for a kinaesthetic semantics centred on topological cognition. His latest book is *Il valore della scienza e questioni affini* (2019).

Giuseppe Samo is Associate Professor of the Department of Linguistics at the Beijing Language and Culture University. He earned his doctorate from the University of Geneva, working on a cartographic approach to Verb Second languages. His current research investigates theoretical predictions adopting the tools and methods of Quantitative Computational Syntax.

Leonardo M. Savoia is Professor Emeritus at the Florence University, where he taught General Linguistics. He is "Corresponding Academician" of the Accademia della Crusca and was President of the Italian Linguistics Society (2003-2007). His research regards phonology and morphosyntax of Italian dialects, Albanian and Aromanian (recently, with M. Rita Manzini, *The Morphosyntax of Albanian and Aromanian Varieties*, De Gruyter Mouton, 2018), sociolinguistics, and, with Benedetta Baldi, communication and political language. Benedetta Baldi and Leonardo M. Savoia have published articles and books, and recently *Linguistica per insegnare*, Zanichelli 2018.

Tesfay Tewolde Yohannes taught linguistics in African universities. Currently he works in collaboration with them. He teaches at the University of Florence. Among his publications, he authored *A Modern Grammar of Tigrinya* (2002), *DPs, Phi-features and Tense in the Context of Abyssinian Semitic Languages* (2016).

Romina Vergari is Lecturer of Hebrew Language at the University of Florence. She has been postdoctoral researcher at the Universities of Haifa and Strasbourg and visiting fellow at the Oxford Centre for Hebrew and Jewish Studies. She is a member of the editorial staff of the Historical and Theological Lexicon of the Septuagint.

Dr Aiqing Wang is Lecturer at the Department of Modern Languages and Cultures, University of Liverpool. After graduating with an MA in Linguistics from UCL, she received a PhD in Linguistics from the University of York. Her doctoral project investigated Late Archaic Chinese syntax.

DIPARTIMENTO DI FORMAZIONE, LINGUE, INTERCULTURA,
LETTERATURE E PSICOLOGIA
BIBLIOTECA DI STUDI DI FILOLOGIA MODERNA: COLLANA, RIVISTE E LABORATORIO

Opere pubblicate

I titoli qui elencati sono stati proposti alla Firenze University Press dal Coordinamento editoriale del Dipartimento di Formazione, Lingue, Intercultura, Letterature e Psicologia e prodotti dal suo Laboratorio editoriale Open Access

Volumi ad accesso aperto
(<http://www.fupress.com/comitatoscientifico/biblioteca-di-studi-di-filologia-moderna/23>)

Stefania Pavan, *Lezioni di poesia. Iosif Brodskij e la cultura classica: il mito, la letteratura, la filosofia*, 2006 (Biblioteca di Studi di Filologia Moderna; 1)
Rita Svandrlik (a cura di), *Elfriede Jelinek. Una prosa altra, un altro teatro*, 2008 (Biblioteca di Studi di Filologia Moderna; 2)
Ornella De Zordo (a cura di), *Saggi di anglistica e americanistica. Temi e prospettive di ricerca*, 2008 (Strumenti per la didattica e la ricerca; 66)
Fiorenzo Fantaccini, *W.B. Yeats e la cultura italiana*, 2009 (Biblioteca di Studi di Filologia Moderna; 3)
Arianna Antonielli, *William Blake e William Butler Yeats. Sistemi simbolici e costruzioni poetiche*, 2009 (Biblioteca di Studi di Filologia Moderna; 4)
Marco Di Manno, *Tra sensi e spirito. La concezione della musica e la rappresentazione del musicista nella letteratura tedesca alle soglie del Romanticismo*, 2009 (Biblioteca di Studi di Filologia Moderna; 5)
Maria Chiara Mocali, *Testo. Dialogo. Traduzione. Per una analisi del tedesco tra codici e varietà*, 2009 (Biblioteca di Studi di Filologia Moderna; 6)
Ornella De Zordo (a cura di), *Saggi di anglistica e americanistica. Ricerche in corso*, 2009 (Strumenti per la didattica e la ricerca; 95)
Stefania Pavan (a cura di), *Gli anni Sessanta a Leningrado. Luci e ombre di una* Belle Époque, 2009 (Biblioteca di Studi di Filologia Moderna; 7)
Roberta Carnevale, *Il corpo nell'opera di Georg Büchner. Büchner e i filosofi materialisti dell'Illuminismo francese*, 2009 (Biblioteca di Studi di Filologia Moderna; 8)
Mario Materassi, *Go Southwest, Old Man. Note di un viaggio letterario, e non*, 2009 (Biblioteca di Studi di Filologia Moderna; 9)
Ornella De Zordo, Fiorenzo Fantaccini (a cura di), *altri canoni / canoni altri. pluralismo e studi letterari*, 2011 (Biblioteca di Studi di Filologia Moderna; 10)
Claudia Vitale, *Das literarische Gesicht im Werk Heinrich von Kleists und Franz Kafkas*, 2011 (Biblioteca di Studi di Filologia Moderna; 11)
Mattia Di Taranto, *L'arte del libro in Germania fra Otto e Novecento: Editoria bibliofilica, arti figurative e avanguardia letteraria negli anni della* Jahrhundertwende, 2011 (Biblioteca di Studi di Filologia Moderna; 12)
Vania Fattorini (a cura di), *Caroline Schlegel-Schelling: «Ero seduta qui a scrivere».* Lettere, 2012 (Biblioteca di Studi di Filologia Moderna; 13)
Anne Tamm, *Scalar Verb Classes. Scalarity, Thematic Roles, and Arguments in the Estonian Aspectual Lexicon*, 2012 (Biblioteca di Studi di Filologia Moderna; 14)

Beatrice Töttössy (a cura di), *Fonti di Weltliteratur. Ungheria*, 2012 (Strumenti per la didattica e la ricerca; 143)

Beatrice Töttössy, *Ungheria 1945-2002. La dimensione letteraria*, 2012 (Biblioteca di Studi di Filologia Moderna; 15)

Diana Battisti, *Estetica della dissonanza e filosofia del doppio: Carlo Dossi e Jean Paul*, 2012 (Biblioteca di Studi di Filologia Moderna; 16)

Fiorenzo Fantaccini, Ornella De Zordo (a cura di), *Saggi di anglistica e americanistica. Percorsi di ricerca*, 2012 (Strumenti per la didattica e la ricerca; 144)

Martha L. Canfield (a cura di), *Perù frontiera del mondo. Eielson e Vargas Llosa: dalle radici all'impegno cosmopolita = Perù frontera del mundo. Eielson y Vargas Llosa: de las raíces al compromiso cosmopolita*, 2013 (Biblioteca di Studi di Filologia Moderna; 17)

Gaetano Prampolini, Annamaria Pinazzi (eds), *The Shade of the Saguaro / La sombra del saguaro: essays on the Literary Cultures of the American Southwest / Ensayos sobre las culturas literarias del suroeste norteamericano*, 2013 (Biblioteca di Studi di Filologia Moderna; 18)

Ioana Both, Ayşe Saraçgil, Angela Tarantino (a cura di), *Storia, identità e canoni letterari*, 2013 (Strumenti per la didattica e la ricerca; 152)

Valentina Vannucci, *Letture anticanoniche della biofiction, dentro e fuori la metafinzione*, 2014 (Biblioteca di Studi di Filologia Moderna; 19)

Serena Alcione, *Wackenroder e Reichardt. Musica e letteratura nel primo Romanticismo tedesco*, 2014 (Biblioteca di Studi di Filologia Moderna; 20)

Lorenzo Orlandini, The relentless body. *L'impossibile elisione del corpo in Samuel Beckett e la* noluntas *schopenhaueriana*, 2014 (Biblioteca di Studi di Filologia Moderna; 21)

Carolina Gepponi (a cura di), *Un carteggio di Margherita Guidacci*, 2014 (Biblioteca di Studi di Filologia Moderna; 22)

Valentina Milli, *«Truth is an odd number». La narrativa di Flann O'Brien e il fantastico*, 2014 (Biblioteca di Studi di Filologia Moderna; 23)

Diego Salvadori, *Il giardino riflesso. L'erbario di Luigi Meneghello*, 2015 (Biblioteca di Studi di Filologia Moderna; 24)

Sabrina Ballestracci, Serena Grazzini (a cura di), *Punti di vista - Punti di contatto. Studi di letteratura e linguistica tedesca*, 2015 (Biblioteca di Studi di Filologia Moderna; 25)

Massimo Ciaravolo, Sara Culeddu, Andrea Meregalli, Camilla Storskog (a cura di), *Forme di narrazione autobiografica nelle letterature scandinave*. Forms of Autobiographical Narration in Scandinavian Literature, 2015 (Biblioteca di Studi di Filologia Moderna; 26)

Lena Dal Pozzo, *New information subjects in L2 acquisition: evidence from Italian and Finnish,* 2015 (Biblioteca di Studi di Filologia Moderna; 27)

Sara Lombardi (a cura di), *Lettere di Margherita Guidacci a Mladen Machiedo*, 2015 (Biblioteca di Studi di Filologia Moderna; 28)

Giuliano Lozzi, *Margarete Susman e i saggi sul femminile*, 2015 (Biblioteca di Studi di Filologia Moderna; 29)

Ilaria Natali, *«Remov'd from Human Eyes»: Madness and Poetry. 1676-1774*, 2016 (Biblioteca di Studi di Filologia Moderna; 30)

Antonio Civardi, *Linguistic Variation Issues: Case and Agreement in Northern Russian Participial Constructions*, 2016 (Biblioteca di Studi di Filologia Moderna; 31)

Tesfay Tewolde, *DPs, Phi-features and Tense in the Context of Abyssinian (Eritrean and Ethiopian) Semitic Languages* (Biblioteca di Studi di Filologia Moderna; 32)

Arianna Antonielli, Mark Nixon (eds), *Edwin John Ellis's and William Butler Yeats's* The Works of William Blake: Poetic, Symbolic and Critical. *A Manuscript Edition, with Critical Analysis*, 2016 (Biblioteca di Studi di Filologia Moderna; 33)

Augusta Brettoni, Ernestina Pellegrini, Sandro Piazzesi, Diego Salvadori (a cura di), *Per Enza Biagini*, 2016 (Biblioteca di Studi di Filologia Moderna; 34)

Silvano Boscherini, *Parole e cose: raccolta di scritti minori*, a cura di Innocenzo Mazzini, Antonella Ciabatti, Giovanni Volante, 2016 (Biblioteca di Studi di Filologia Moderna; 35)

Ayşe Saraçgil, Letizia Vezzosi (a cura di), *Lingue, letterature e culture migranti*, 2016 (Strumenti per la didattica e la ricerca; 183)

Michela Graziani (a cura di), *Trasparenze ed epifanie. Quando la luce diventa letteratura, arte, storia, scienza*, 2016 (Biblioteca di Studi di Filologia Moderna; 36)

Caterina Toschi, *Dalla pagina alla parete. Tipografia futurista e fotomontaggio dada*, 2017 (Biblioteca di Studi di Filologia Moderna; 37)

Diego Salvadori, *Luigi Meneghello. La biosfera e il racconto*, 2017 (Biblioteca di Studi di Filologia Moderna; 38)

Sabrina Ballestracci, *Teoria e ricerca sull'apprendimento del tedesco L2*, 2017 (Strumenti per la didattica e la ricerca; 194)

Michela Landi (a cura di), *La double séance. La musique sur la scène théâtrale et littéraire / La musica sulla scena teatrale e letteraria*, 2017 (Biblioteca di Studi di Filologia Moderna; 39)

Fulvio Bertuccelli (a cura di), *Soggettività, identità nazionale, memorie. Biografie e autobiografie nella Turchia contemporanea*, 2017 (Biblioteca di Studi di Filologia Moderna; 40)

Susanne Stockle, *Mare, fiume, ruscello. Acqua e musica nella cultura romantica*, 2018 (Biblioteca di Studi di Filologia Moderna; 41)

Gian Luca Caprili, *Inquietudine spettrale. Gli uccelli nella concezione poetica di Jacob Grimm*, 2018 (Biblioteca di Studi di Filologia Moderna; 42)

Dario Collini (a cura di), *Lettere a Oreste Macrì. Schedatura e regesto di un fondo, con un'appendice di testi epistolari inediti*, 2018 (Biblioteca di Studi di Filologia Moderna; 43)

Simone Rebora, *History/Histoire e Digital Humanities. La nascita della storiografia letteraria italiana fuori d'Italia*, 2018 (Biblioteca di Studi di Filologia Moderna; 44)

Marco Meli (a cura di), *Le norme stabilite e infrante. Saggi italo-tedeschi in prospettiva linguistica, letteraria e interculturale*, 2018 (Strumenti per la didattica e la ricerca; 203)

Francesca Di Meglio, *Una muchedumbre o nada: Coordenadas temáticas en la obra poética de Josefina Plá*, 2018 (Biblioteca di Studi di Filologia Moderna; 45)

Barbara Innocenti, *Il piccolo Pantheon. I grandi autori in scena sul teatro francese tra Settecento e Ottocento*, 2018 (Biblioteca di Studi di Filologia Moderna; 46)

Oreste Macrí, Giacinto Spagnoletti, *«Si risponde lavorando». Lettere 1941-1992, a cura di Andrea Giusti*, 2019 (Biblioteca di Studi di Filologia Moderna; 47)

Michela Landi, *Baudelaire et Wagner*, 2019 (Biblioteca di Studi di Filologia Moderna; 48)

Sabrina Ballestracci, *Connettivi tedeschi e poeticità: l'attivazione dell'interprete tra forma e funzione. Studio teorico e analisi di un caso esemplare*, 2019 (Biblioteca di Studi di Filologia Moderna; 49)

Fiorenzo Fantaccini, Raffaella Leproni (a cura di), *"Still Blundering into Sense". Maria Edgeworth, her context, her legacy*, 2019 (Biblioteca di Studi di Filologia Moderna; 50)

Arianna Antonielli, Donatella Pallotti (a cura di), *"Granito e arcobaleno". Forme e modi della scrittura auto/biografica*, 2019 (Biblioteca di Studi di Filologia Moderna; 51)

Francesca Valdinoci, *Scarti, tracce e frammenti: controarchivio e memoria dell'umano*, 2019 (Biblioteca di Studi di Filologia Moderna; 52)

Sara Congregati (a cura di), *La* Götterlehre *di Karl Philipp Moritz. Nell'officina del linguaggio mitopoietico degli antichi*, traduzione integrale, introduzione e note di Sara Congregati, 2020 (Biblioteca di Studi di Filologia Moderna; 53)

Gabriele Bacherini, *Frammenti di massificazione: le neoavanguardie anglo-germanofone, il cut-up di Burroughs e la pop art negli anni Sessanta e Settanta*, 2020 (Biblioteca di Studi di Filologia Moderna; 54)

Inmaculada Solís García y Francisco Matte Bon, *Introducción a la gramática metaoperacional*, 2020 (Strumenti per la didattica e la ricerca; 216)

Barbara Innocenti, Marco Lombardi, Josiane Tourres (a cura di), *In viaggio per il Congresso di Vienna: lettere di Daniello Berlinghieri a Anna Martini, con un percorso tra le fonti archivistiche in appendice*, 2020 (Biblioteca di Studi di Filologia Moderna; 55)

Elisabetta Bacchereti, Federico Fastelli, Diego Salvadori (a cura di), *Il* graphic novel. *Un crossover per la modernità*, 2020 (Biblioteca di Studi di Filologia Moderna; 56)

Tina Maraucci, *Leggere Istanbul. Memoria e lingua nella narrativa turca contemporanea*, 2020 (Biblioteca di Studi di Filologia Moderna; 57)

Valentina Fiume, *Codici dell'anima: itinerari tra mistica, filosofia e poesia. Con un'antologia di scritti al femminil*e, 2021 (Biblioteca di Studi di Filologia Moderna; 58)

Riviste ad accesso aperto
(<http://www.fupress.com/riviste>)

«Journal of Early Modern Studies», ISSN: 2279-7149

«LEA - Lingue e Letterature d'Oriente e d'Occidente», ISSN: 1824-484X

«Quaderni di Linguistica e Studi Orientali / Working Papers in Linguistics and Oriental Studies», ISSN: 2421-7220

«Studi Irlandesi. A Journal of Irish Studies», ISSN: 2239-3978