

Studi e Saggi Linguistici

Direzione Scientifica / Editor in Chief

Giovanna Marotta, *Università di Pisa*

Comitato Scientifico / Advisory Board

Béla Adamik, *University of Budapest*
Michela Cennamo, *Università di Napoli «Federico II»*
Bridget Drinka, *University of Texas at San Antonio*
Giovambattista Galdi, *University of Gent*
Nicola Grandi, *Università di Bologna*
Adam Ledgeway, *University of Cambridge*
Luca Lorenzetti, *Università della Tuscia*
Elisabetta Magni, *Università di Bologna*
Patrizia Sorianello, *Università di Bari*
Mario Squartini, *Università di Torino*

Comitato Editoriale / Editorial Board

Marina Benedetti, *Università per Stranieri di Siena*
Franco Fanciullo, *Università di Pisa*
Marco Mancini, *Università di Roma «La Sapienza»*

Segreteria di Redazione / Editorial Assistants

Rosangela Lai	<i>e-mail: rosangela.lai@unipi.it</i>
Francesco Rovai	<i>e-mail: francesco.rovai@unipi.it</i>
Lucia Tamponi	<i>e-mail: lucia.tamponi@fileli.unipi.it</i>

I contributi pervenuti sono sottoposti alla valutazione di due revisori anonimi.

All submissions are double-blind peer reviewed by two referees.

Studi e Saggi Linguistici è indicizzato in / *Studi e Saggi Linguistici* is indexed in

ERIH PLUS (European Reference Index for the Humanities and Social Sciences)
Emerging Sources Citation Index - Thomson Reuters
L'Année philologique
Linguistic Bibliography
MLA (Modern Language Association Database)
Scopus

STUDI E SAGGI LINGUISTICI

LXI (1) 2023

rivista fondata da
TRISTANO BOLELLI



Edizioni ETS



STUDIE SAGGI LINGUISTICI

www.studiesaggilinguistici.it

SSL electronic version is now available with OJS (Open Journal Systems)

Web access and archive access are granted to all registered subscribers

Abbonamento, compresa spedizione
individuale, Italia € 50,00
individuale, Estero € 70,00
istituzionale, Italia € 60,00
istituzionale, Estero € 80,00
Bonifico su c/c Edizioni ETS srl
IBAN IT 21 U 03069 14010 100000001781
BIC BCITITMM
Causale: Abbonamento SSL

Subscription, incl. shipping
individual, Italy € 50,00
individual, Abroad € 70,00
institutional, Italy € 60,00
institutional, Abroad € 80,00
Bank transfer to Edizioni ETS srl
IBAN IT 21 U 03069 14010 100000001781
BIC BCITITMM
Reason: Subscription SSL

L'editore non garantisce la pubblicazione prima di sei mesi dalla consegna in forma definitiva di ogni contributo.

Registrazione Tribunale di Pisa 12/2007 in data 20 Marzo 2007

Periodicità semestrale

Direttore responsabile: Alessandra Borghini

ISBN 978-884676674-8

ISSN 0085 6827

RISERVATO OGNI DIRITTO DI PROPRIETÀ E DI TRADUZIONE



Indice

Saggi

Similar forms with different functions. The (un)learning of the Italian periphrasis <i>andare a</i> + infinitive by L1-Spanish speakers PAOLO DELLA PUTTA, FRANCESCA STRIK-LIEVERS	9
A correlation between applicative marker types and word order patterns: diachronic perspective DEOKHYUN NAM	51
Armenian stem classes and the Western-Middle Iranian oblique plural case MARCO FATTORI	107
Sulla polifunzionalità di <i>vienas</i> “uno” in lituano antico ADRIANO CERRI	123
Il trattamento di /w/ in posizione iniziale in italiano è influenzato dall’ortografia? Indizi dalla selezione dell’articolo maschile PIERO COSSU	161

SL

Saggi

Similar forms with different functions. The (un)learning of the Italian periphrasis *andare a* + infinitive by L1-Spanish speakers

PAOLO DELLA PUTTA, FRANCESCA STRIK-LIEVERS

ABSTRACT

Despite their formal similarity, the Italian periphrasis *andare a* “go to” + infinitive and the Spanish periphrasis *ir a* “go to” + infinitive display different values. In this paper, we investigate how Spanish (L1) speakers acquire *andare a* + infinitive when learning Italian as a second language (L2). This case study offers a good vantage point from which to observe the interaction between learning (values that are present in L2 but not in L1) and unlearning (values that are present in L1 but not in L2) in the acquisition of L2 forms that are structurally but not functionally similar to L1 forms. An experimental study based on acceptability judgements shows that while a long and intense exposure to the L2 input allows both learning and unlearning, L2 instruction not focused on the two periphrases and with little exposure to input is less successful: unlearning of non-target values is only reached by advanced speakers, who moreover generalise the inhibition of the L1-transfer to values that are accepted by native speakers. This suggests that in cases like the one under investigation here a focused pedagogical intervention may be useful.

KEYWORDS: Romance periphrases, aspect, unlearning, second language acquisition.

1. Introduction

Learners of a second language (L2) that is closely related¹ to their first language (L1) enjoy many advantages in their acquisition process but must also face specific challenges. One such challenge consists in learning L2 forms that are very similar to L1 forms but have completely

¹ In this paper, we operationalise the closeness of two languages as a function resulting from their belonging to the same language family and their mutual intelligibility. Following this operationalisation, GOOSKENS *et al.* (2018) demonstrated that, within the Romance family, Italian and Spanish is the most closely related language pairing after Spanish and Portuguese.

or partly different meanings and functions or are used in different varieties and contexts in the two languages. This is the case for ‘forms’ at all levels of analysis. For example, Spanish and Italian both have the grapheme <d>, but while in Italian <d> always corresponds to [d], in Spanish it corresponds to either [d] or [ð] depending on the phonetic context. At a lexical level, closely related languages often have so-called ‘false friends’, that is, words with similar or identical form but with different meanings, such as *palestra*: Portuguese “conference”, Italian “gym”. In these cases of form and function mismatch, learners need not only to *learn* new functions to be added to the interlanguage, but also to *unlearn* (Gass and Mackey, 2002) functions that are associated with a similar form in the L1 but not in the L2. In our examples, L1-Italian speakers must not only learn that in Spanish some <d>s are to be read as [ð], and that Portuguese *palestra* means “conference”, but they also have to inhibit the L1 transfer that would lead them to read all <d>s as [d] and to attribute to *palestra* the meaning “gym”.

Italian and Spanish being very closely related, L2 acquisition in both directions abounds with such cases of mismatch. Among them are many verbal periphrases, which are present in the two languages and are often formally similar (Ambrosini and Della Putta, 2021). This is the case of the periphrasis which is the subject of this study, that is, Italian *andare a* “go to” + infinitive. Despite the formal correspondence between *andare a* + infinitive and its Spanish counterpart *ir a* “go to” + infinitive, the two periphrases display different values. The Italian structure has a culminative and, more rarely, a prospective aspectual value, while the Spanish one is widely used with a temporal value, and also has aspectual and modal values (Olbertz, 1998; Squartini, 1998; Garachana, 2018; see § 2 for a more detailed description). Moreover, the Italian periphrasis is less frequently employed overall, as some of its uses tend to be restricted to specific diastratic and diaphasic varieties, and it is less codified (e.g., it is often absent from descriptive and pedagogical grammars) compared to its Spanish counterpart (see Renzi, 2012: 103).

The question we address in this paper is: how do L1-Spanish speakers of Italian acquire the periphrasis *andare a* + infinitive? To

answer this question, we conducted an experimental study using acceptability judgements. Three groups of informants participated in the study: (i) L1-Spanish speakers with different levels of proficiency in Italian, who were formally instructed in Italian in their home country, (ii) L1-Spanish speakers of Italian with no formal instruction in the language, who have lived in Italy for longer than one year, and (iii) a control group of Italian native speakers.

We believe that this case study offers a good vantage point from which to observe the interaction between learning and unlearning in the acquisition of L2 forms that are structurally but not functionally similar to L1 forms. More specifically, we aim to investigate the role played by two factors. First, the level of L2 proficiency: do learners' judgements become increasingly native-speaker-like as proficiency increases, as one would expect? Second, L2 input: is there a difference between learners who are formally instructed in their home country and learners who are not formally instructed but have been living in Italy, and do judgements become increasingly native-like as the length of stay in Italy increases?

2. *Italian andare a + infinitive and Spanish ir a + infinitive*

A large number of studies have focused on verbal periphrases in the Romance languages in general, and on verbal periphrases in Italian and Spanish more particularly (see, among many others, Dietrich, 1973; Olbertz, 1998; Squartini, 1998; Gómez Torrego, 1999; Bertinetto, 1991; RAE, 2009: 529-556; Fábregas, 2019; Laca, 2021). Studies specifically dedicated to *andare a* + infinitive (Amenta and Strudsholm, 2002; Valentini, 2007) and *ir a* + infinitive (Fernández De Castro, 1999; Matte Bon, 2006; Bravo, 2008; Bravo and Laca, 2011) are also numerous. The abundance of research on the topic is at least partly due to the fact that on many even quite fundamental issues, scholars are far from having reached a definitive position.

First, criteria for defining what counts as a periphrasis are still highly debated as not all of these work equally well across languages,

or across different sets of periphrases in a single language, to the point that periphrasticity is often considered as a gradient notion (Laca, 2021). As regards the Italian periphrasis we are focusing on here, it seems that it does indeed have uses that can be considered periphrastic according to the main criteria of periphrasticity (Bertinetto, 1990). For instance, the verb *andare* is largely desemantised, the relative order of *andare* and the infinitive is fixed, and it is only possible to insert 'light' elements between the two (see Amenta and Strudsholm, 2002; Valentini, 2007). Other criteria that are commonly used to define periphrasticity seem to be less relevant for *andare a* + infinitive: clitic climbing, for example, also applies when *andare a* + infinitive is not used periphrastically (Strik-Lievers, 2017: 171).

Second, both the Italian and the Spanish periphrasis co-exist with synthetic forms that have partly coinciding functions. Myriad papers have been dedicated, for example, to the competition between *ir a* + infinitive and the synthetic future, and have tried to disentangle the peculiarities of the two forms (among others, Matte Bon, 2006; Bravo, 2008; Lara-Bermejo, 2016). The problem is further complicated by the fact that the distinction between the periphrastic and the synthetic future is not of a semantic nature only: variation in usage is often determined on diatopic, diaphasic, diastratic, or pragmatic grounds. As observed by Laca (2021), this complex situation, where the competition between synthetic and analytic forms is played out on several fronts, increases the difficulty in describing the features of periphrases. Synchronic dimensions of variation are also crucial for *andare a* + infinitive, as its periphrastic uses tend to be limited to specific varieties of Italian, being virtually absent (and often stigmatised) in others. In Strik-Lievers (in press) it is shown that periphrastic *andare a* + infinitive is more common in written language, while in spoken language it is very rare, except in expository discourse (e.g., university lectures). Its variability and marginality exclude *andare a* + infinitive from L1 and, more importantly for our purposes, L2 pedagogical grammars (Oricco, 2020).

The extent and variety of the literature existing for both the Italian and the Spanish periphrasis does not allow us to do it justice here. What is needed for our purposes is a contrastive outline, based on this

literature, of the *main* values of the two periphrases. This will help us in telling apart the cases in which learning is required (i.e., when the Italian periphrasis has values that are absent in the Spanish) from those in which unlearning is required (i.e., when the Italian periphrasis does not have values that are present in the Spanish).

Bearing in mind the many pitfalls involved in the description of the two periphrases, we are aware that the categorisation proposed below in § 2.1 is far from complete and indisputable, and that the separation between different values as well as the interpretation of individual examples may be more nuanced than presented here. Some rigidity in categorisation is however required for operational reasons, given the experimental nature of this study. The decision to ask Italian native speakers for scalar acceptability judgements rather than binary ones is motivated precisely by an attempt to mitigate this problem (see § 4.1). For instance, the fact that a given sentence including *andare a* + infinitive is commonly used in some varieties but almost never in others might be reflected by a dubitative intermediate rating. Through scalar acceptability judgements, in fact, a ‘gradient of acceptability’ can be constructed, and therefore do justice to how speakers perceive the correctness of a sentence, which ultimately depends on a complex interplay of different linguistic constraints (Francis, 2022: 1-16). After this premise, we can now contrastively illustrate the main values displayed by the Italian periphrasis and its Spanish counterpart.

2.1. *Uses of the two periphrases: a contrastive overview*

In addition to their periphrastic uses, both *andare a* + infinitive and *ir a* + infinitive also display non-periphrastic, fully compositional uses, where *andare* and *ir* keep their physical motion meaning, as in (1) and (2):

- (1) *Ora vado a comprare le ciliegie al mercato.*²
 now I.go to buy the cherries to.the market
 “I now go to buy cherries at the market.”

² All Italian examples are our own.

- (2) *No pude ir a verla.*
 not I.could go to see.her
 “I could not go to see her.”

(RAE, 2009: 2158)

In what follows, we only focus on: (i) periphrastic uses, because the non-periphrastic uses exemplified above do not present significant differences between the two languages nor specific challenges for learners, and (ii) uses that are displayed when *andare* and *ir* are in their present indicative, since this is the only form we use in our test sentences (see § 5). We first illustrate the uses that characterise the Italian periphrasis but not the Spanish, and then those that characterise the Spanish periphrasis but not the Italian.

2.1.1. *Values of Italian andare a + infinitive*

Andare a + infinitive is used with a culminative (i) and, to a lesser extent, a prospective aspectual value (ii). In some cases, the semantic contribution is harder to classify, as it is rather bleached (iii). Additionally, it may form collocations (iv) when specific verbs appear in the infinitive (e.g., *sapere* “to know”, *finire* “to finish”).

(i) *Aspectual (culminative)*

The Italian periphrasis mainly has an aspectual value that can be defined as culminative, following Veland (2004; or resolute, in Bertinetto, 1991). *Andare* maintains here the deictic orientation that characterises its lexical use (Bertinetto, 1991: 141), describing a metaphorical and non-planned motion that culminates in the event expressed by the verb in the infinitive, as in (3):

- (3) *Questo stadio va a sostituire quello vecchio.*
 this stadium goes to replace that old
 “This stadium replaces / ends up replacing the old one.”

(ii) *Aspectual (prospective)*

Andare a + infinitive can also be used to express an action that will occur in the near future, as in (4):

- (4) *Vado a illustrare la mia proposta.*
 I.go to explain the my proposal
 “I’m going to explain my proposal.”

This value is shared by the Spanish periphrasis, which however displays a wider range of prospective aspectual uses. Perhaps more importantly for our purposes, while the prospective use is prominent in Spanish, being widely employed across varieties and clearly codified by grammars, in Italian it is rather marginal and mostly bound to specific varieties. For instance, it is found in the semi-scientific Italian (Sobrero, 2003: 241) of university students’ essays and dissertations, where the prospective use of the periphrasis – condemned by prescriptive publications since the 19th century – has a textual-deictic function (Strik-Lievers, in press). A related use is the one exemplified in (5), where the periphrasis refers to an action that the speaker is about to perform, and is typically found in videos illustrating practical procedures such as recipes (Frosini, 2020):

- (5) *Andiamo a separare l’ albume dal tuorlo.*
 we.go to separate the egg.white from.the yolk
 “Let’s separate the egg white from the yolk.”

(iii) *Pleonastic*

In varieties such as scientific and semi-scientific writing, which often feature ‘empty’ analytic forms (e.g., *viene ad essere* lit. “comes to be” instead of *è* “is”, Berruto, 2012: 100), *andare a* + infinitive is frequently also used with values that do not easily fall into the categories described above, or into other categories. The aspectual content of the periphrasis is rather bleached, to the point that it seems to be almost pleonastic and can be replaced by a synthetic form with little if any semantic loss (Renzi, 2019: 15-16). For example, in (6), while possibly having a culminative nuance, *va a indicare* could be replaced by the present indicative *indica* without semantic loss.

- (6) *L' alto numero di morti va a indicare*
 the high number of dead goes to indicate
un peggioramento della situazione.
 a worsening of the situation
 "The high number of deaths indicates a worsening of the situation."

Valentini (2007) observes that in this case the role of the periphrasis is mainly to highlight the functionality of the event in the predicate, that is, in (6) *va a indicare* could be paraphrased as "has the function of indicating".

Example (7) is possibly more connected to the prospective use, but again replacing *va a diminuire* with the present indicative *diminuisce* does not involve appreciable semantic differences.

- (7) *Piano piano va a diminuire il numero*
 gradually goes to decrease the number
di contagi.
 of infections
 "The number of infections is gradually decreasing."

(iv) *Collocations*

Finally, the periphrasis displays some uses that can be considered collocational, being characterised by little to no productivity and a meaning that cannot be included in the aspectual categories above. For instance, the imperative *vai a sapere/capire* ("to know/understand") is used as a discourse marker, as in (8):

- (8) *Potresti vincere, vai a sapere.*
 you.might win go to know
 "You might win, who knows."

2.1.2. *Values of Spanish ir a + infinitive*

Ir a + infinitive is frequently used, and it displays various temporal (i, ii), aspectual (iii), and modal (iv) values. Additional less common values have also been identified in the literature (v).

(i) *Temporal (future)*

The main periphrastic use of *ir a* + infinitive is temporal (RAE, 2009: 2113). Here, the periphrasis expresses future tense, and in Spanish it is more and more frequently replacing the synthetic future, as in (9) (Matte Bon, 2006; Bravo, 2008; Lara-Bermejo, 2016):

- (9) *Si te lo digo, ¿me vas a querer más?*
 if you it say me you.go to love more
 “If I tell it to you, will you love me more?”

(Gómez Torrego, 1999: 3368)

(ii) *Temporal (planned future)*

The temporal periphrasis can also take on modal nuances of intentionality, thus expressing what is often called planned future, as in (10) where – unlike in the case of (9) – the periphrasis cannot be replaced by the synthetic future:

- (10) *Si vas a venir, dínoslo.*
 if you.go to come tell.us.it
 “If you are going to come, do tell us.”

(Gómez Torrego, 1999: 3369)

(iii) *Aspectual (prospective / immediate prospective)*

In this use, *ir a* + infinitive refers to an action due to start in the near future. It can refer to both prospective aspect and immediate prospective aspect (see Dik, 1997: 238-239 for this distinction, exemplified in English by *John is going to cry* and *John is about to cry* respectively). While the prospective value is shared by the Italian periphrasis, the immediate prospective value (also labeled ‘imminent’), exemplified in (11), is not.

- (11) *Señores, va a comenzar el partido.*
 gentlemen goes to begin the match
 “Gentlemen, the match is about to start.”

(Gómez Torrego, 1999: 3368)

Incidentally, it is to be noted that the distinction between the aspectual uses and the (increasingly expanding) temporal uses of the periphrasis is much discussed in the literature (see Bravo, 2008: Ch. 3) and is not always straightforward, with many examples that are open to both interpretations. For example, *va a llover* (Bravo, 2008: 201) can be interpreted both temporally (“it will rain”, e.g., in weather forecast for next week) and aspectually (“it is about to rain”, e.g., spotting dark clouds in the sky and hearing thunder).

(iv) *Modal*

The Spanish periphrasis can also display a modal value of possibility or probability which is close to that of conjectural future (Bravo, 2008: 22), as in (12), although this is less frequently attested compared to the temporal and aspectual use.

- (12) *Va a haberlo matado Juan.*
 goes to have.it killed Juan
 ‘‘It was probably Juan who killed him.’’

(Gómez Torrego, 1999: 3370)

(v) *Other values*

Finally, *ir a* + infinitive has several uses that do not fall within the categories above and which, unlike Italian ‘pleonastic’ uses, are relatively frequent and can be found across genres and varieties. This is the case with the value that is described as evidential in RAE (2009: 2159)³, that found in exclamative sentences (RAE, 2009: 2116)⁴, and the many values that we did not consider here because they are only available when *ir* is inflected in forms other than the present indicative (for a detailed description of all the

³ An example of evidential use proposed in RAE (2009: 2159) is: *¿Dónde está Juan? ¿Dónde va a estar? En el bar* ‘Where is Juan? Where will he be? At the bar’: the fact that Juan is at the bar, at the time of the utterance, is presented as evident.

⁴ Among the examples of periphrasis in exclamative sentences, RAE (2009: 2116) has: *¿Cómo le vas a decir eso a un amigo!* ‘How can you say that to a friend!’.

uses, see Bravo, 2008; RAE, 2009: 529-556; Gómez Torrego, 1999). Collocational uses, some of which closely correspond to the Italian ones, are also attested⁵.

3. *L2 learning of similar structures with different functions*

The interlanguage resulting from the acquisitional process of a L2 closely related to the L1 presents, from its very beginning, a peculiar morphosyntactic organisation, usually more complex than the one found in the interlanguage of learners with an L1 distant from the L2 (Díaz *et al.*, 2007; Ringbom, 2007: 1-4; Ringbom and Jarvis, 2009). This phenomenon is easily observed in the productions of L1-Spanish speakers learning Italian. Facilitated by the structural proximity of the two codes and by a generally rather *unconcerned* psychotypological attitude (Kellerman, 1983), the ‘hispanohablantes’ attempt – early on – to produce complex grammatical categories with complex morphosyntactic means, which however are often not fully target-like (Schmid, 1994: 196). A detailed analysis of their interlanguage reveals that it is characterised by both simplifications and complexifications (Schmid, 1994: 208) of the Italian morphosyntax. The former phenomenon encompasses, for example, paradigmatic regularisation and allomorphy reduction, which also concerns learners of distant languages; the latter phenomenon is characterised by the recourse to structures or functions typical of Spanish that are erroneously transferred to Italian, so much so as to obtain an interlanguage that is morphosyntactically more complex, in some areas, than Italian itself (Schmid, 1994: 210; Della Putta, 2016).

The transfer of L1 features that are not target in the L2 occurs due to an excessive reliance on the strategy of congruence (Schmid, 1994) between the two languages, which gives L1-Spanish speakers the belief that they can express, in Italian, even complex functions through

⁵ For instance, *vete a saber* functions as a discourse marker, much like Italian *vai a sapere*: *Vete a saber lo que habrían dicho de mí!* ‘Who knows what they will have said about me!’ (GÓMEZ TORREGO, 1999: 3372).

morphosyntactic realisations typical of their L1. According to Corder (1984), this is motivated by the fact that learners of closely related languages base their interlinguistic hypotheses on the characteristics of the L1, gradually adapting and modifying them on the basis of the data coming from the L2 input. Learners are thus engaged in a process of falsification and/or confirmation of their hypotheses, that are *prima facie* based on the L1 model. This is called a restructuring *continuum*: the learner, through exposure to the input and its analysis, refutes or reinforces the ‘instinctive’ hypotheses of congruence between the two languages.

To avoid the erroneous transfer of L1-based values, L1-Spanish speakers of Italian need to unlearn such values. The term ‘unlearning’ refers to a cognitive process that enables the individual to inhibit past knowledge and behavioural routines that have become obsolete and may undermine new knowledge acquisition, therefore affecting the adaptation to environmental change (see Grisold and Kaiser, 2017). In L2 speakers, the same inhibitory processes are activated to avoid code-mixing and L1 transfer during language production and comprehension (Zirnstein *et al.*, 2018; McManus, 2022: 40-44). L2 speakers need, in fact, to inhibit L1 processing habits that may lead to L1-L2 conflict and overlapping. Therefore, while acquiring an L2, individuals do not unlearn an L1 property or structure in the sense that they forget it, but, rather, they inhibit its activation in L2 contexts that potentially but wrongly trigger it. This is a difficult issue, as learners must notice the presence of an absence, that is, the non-appearance of an L1 form in the L2 or of an L1 function or meaning that is not mapped on an L2 form. Unlearning is therefore a process driven by indirect negative evidence, i.e., by the absence of something in the L2 input (Gass and Mackey, 2002; Pearl and Mis, 2016; Schwartz and Goad, 2017). The process of learning, by contrast, is driven by positive evidence, i.e., the appearance, albeit with varying degrees of salience and frequency, of a new form or of new functions or meanings in the L2 input (DeKeyser, 2016).

Many studies put forward the idea that unlearning is a more demanding task than learning (Schwartz and Goad, 2017: 237;

4. *L2 learning of andare a + infinitive by Spanish speakers*

(13) *Da domani vado a essere un bravo studente.*
 from tomorrow I.go to be a good student
 (Della Putta, 2016: 243)

⁶ In a cross-sectional study, DELLA PUTTA and STRIK-LIEVERS (2020) looked at how L1-Spanish students of Italian unlearned the aspectual value of durativity of the periphrasis *estar* + gerund, formally similar to Italian *stare* + gerund, on which only the value of progressivity is mapped. The results showed that the unlearning of the use of *stare* + gerund with durative aspectual values improves alongside the general level of proficiency of the learners but nevertheless, even at advanced levels, it reached only 57% correctness: a certain degree of uncertainty remains.

The acquisition of *andare a* + infinitive by L1-Spanish speakers is particularly complex because it requires both learning and unlearning. On the one hand, Spanish values which are not target in Italian have to be unlearned. On the other hand, new values, which are target in Italian but are not present in Spanish have to be learned. Against this background, our study seeks to answer the following research questions:

- RQ1 – What happens when learners deal with a L2 periphrasis that sounds familiar, but whose values and uses are different from those of the ‘corresponding’ L1 one?
- RQ2 – Is the learning of ‘new’ periphrastic functions easier or more difficult than the unlearning of L1-functions that are not target in the L2?
- RQ3 – Is the ‘attuning’ to target-like Italian features of *andare a* + infinitive modulated by the proficiency level of the L1-Spanish speakers?
- RQ4 – Is the ‘attuning’ to target-like Italian features of *andare a* + infinitive modulated by the amount of input exposure of the L1-Spanish speakers?

4.1. *Methods*

To answer our research questions, we designed an on-line, untimed test that consists of 48 items (see the *Appendix*), of which 24 are target sentences and 24 are fillers. The 24 target items include both arguably acceptable Italian uses of *andare a* + infinitive (see § 2.1.1) and arguably unacceptable sentences that are calques of uses of *ir a* + infinitive which are theoretically described as grammatical in Spanish but ungrammatical in Italian (see § 2.1.2), as in (14):

- (14) *’Se vai ad andare alla festa, dimmelo!*
 if you.go to go to.the party tell.me

Participants were asked to judge the items on a 5-pointscale (1 being ‘fully incorrect’ and 5 being ‘fully correct’). This allows the

creation of a gradient of acceptability, which is a particularly valuable measure when dealing with rather unstable structures. For example, it is useful for the analysis of structures which show diastatic or diaphasic variation, which are constrained by pragmatic or stylistic factors or display some unstable grammatical features given, e.g., language contact or evolution (Schütze, 2019: 55-96; Francis, 2022: 202-210). This is the case of *andare a* + infinitive which, as discussed in § 2, is a periphrasis whose use is quite marginal and mostly bound to specific varieties, contexts of use and functions. For this reason, we did not opt for a production test, which can often fail to elicit rare, unstable or sociolinguistically constrained structures (Francis, 2022: 4-5). All sentences were in the present tense (*presente indicativo*), except for those with the collocations *vai a sapere* and *vai a capire*, where the imperative in any case shares its form with the present indicative. They were created using simple language, with morphosyntactic phenomena that are usually presented in a beginners' course of L2 Italian; we furthermore tried to use a transparent lexicon, given the high degree of intercomprehension between Spanish and Italian (Donato and Pasquarelli-Gascon, 2015). We administered the test to three groups of speakers: (i) native speakers of Italian; (ii) instructed L1-Spanish speakers of Italian; (iii) non-instructed L1-Spanish speakers of Italian. Native speakers' judgements are used as a baseline against which we evaluate learners' judgements. Given the variability of *andare a* + infinitive and given the lack of a clear and unified description of its values in standard Italian, we felt it was vital to ascertain how its uses were judged by Italian native speakers, and not to simply rely on theoretical descriptions or on our own acceptability judgements.

The instructed L1-Spanish speakers studied Italian in their home country; this group of participants allows us to explore L2 proficiency level as a variable (RQ3). The non-instructed L1-Spanish speakers had not received formal instruction of Italian but had been living in Italy for more than one year. This group of participants allows us to explore the amount of input (operationalised as length of stay in Italy) as a variable (RQ4).

4.2. *Participants*

For the ‘native speakers of Italian’ group, we enrolled in the study 165 Italian speakers (mean age 38.7, range 17-76; 65% F, 35% M). The subjects were randomly chosen: the test was sent to a first batch of 25 acquaintances, whom we asked to extend the test to others such as family members, colleagues, etc. The 165 subjects were mainly born and raised in Northern Italy (77%); 78% of them hold a university degree (bachelor, master, or PhD) and the remaining 22% have a high school diploma. The group is thus particularly homogenous as regards the diatopic variety of Italian used and the educational level.

For the ‘instructed L1-Spanish speakers of Italian’ group, we enrolled 91 students⁷ of L2 Italian (mean age 33.3, range 19-72; 52% F, 48% M). Five subjects were removed from the sample for the lack of clear information about their study course; the remaining 86 students were divided into three competence groups: beginners ($N = 34$), intermediate ($N = 31$) and advanced ($N = 21$). The competence of the subjects was measured on the last test they took at their university or school. In 88% of cases, the answers were given according to the Common European Framework of Reference for Languages (CEFRL) competence levels. In the remaining 12% of cases, the subjects did not evaluate their competence level according to the CEFRL descriptors, but their self-assessment was clear enough for us to place them in one of the six CEFRL competence levels. Furthermore, we asked the subjects to state how long they had been studying Italian, and we found a linear, strong relationship between the two variables ‘competence level’ and ‘study length’ (Pearson $r = 0.71$). In general, beginners studied Italian for less than 2 years, intermediate learners from 2 to 4 and advanced learners for more than 4 years.

⁷ Among these, 13 declared themselves to be bilingual Catalan/Spanish speakers. As Catalan does not make use of a periphrastic structure similar to *andare/ir a* + infinitive to express temporal and aspectual values (LARA-BERMEJO, 2021), and as *ir a* + infinitive is less used in the Spanish spoken in Catalan-speaking regions of Spain (ENRIQUE-ARIAS, 2010), we decided to ascertain if the answers of this subset of informants were somehow different. We therefore run a U-Mann Whitney test with a 2-level independent variable (Catalan or Spanish as L1) and, as dependent variables, the ratings given to the 24 target items. In no cases were significant differences found: it was therefore not necessary to exclude the 13 bilingual subjects from our sample.

It should be noted here that the uses of *andare a* + infinitive are not dealt with in L2 pedagogical grammars (see § 2), and that, more generally, periphrases are not analysed in a contrastive manner in L2 Italian courses for Spanish speakers (Ambrosini and Della Putta, 2021). Although we cannot exclude the possibility that teachers corrected non-target uses of *andare a* + infinitive, we can suppose that the instructed group participants did not receive a robust explicit instruction on the differences between *andare a* + infinitive and *ir a* + infinitive. In addition, these learners had studied Italian in Spain only and had not spent long periods in an Italian-speaking country: they had never had everyday, direct access to standard and nonstandard Italian.

For the ‘non-instructed L1-Spanish speakers of Italian’ group, we enrolled 69 participants (mean age 36.8, range 24-67; 58% F, 42% M) who had been living in Italy for at least one year. All of these were immigrant workers who had received no or nearly no formal Italian instruction. By ‘nearly no instruction’ we mean: (i) that the amount of formal instruction for Italian declared by the subjects did not exceed six months, and (ii) that in these 6 months the instruction did not have a Spanish-Italian contrastive basis, i.e., it was designed for multilingual classes. Self-study, when reported, was accepted. Against this criterion, seven subjects had to be excluded from the original sample as they had received more than 6 months of formal instruction. Out of the remaining 62 individuals, 2 were Spanish citizens, while all the others came from South America⁸, mainly from Chile, Peru, or Bolivia. 48 of the 62 participants lived in the Turin area, 11 lived in Lombardy and the remaining 3 lived in Tuscany, Lazio, and Sicily. The regional Italian they have been exposed to is therefore similar to that of the Italian informants, with a predominance of Northern-Italian varieties. 88.5% of the non-instructed population said they use Italian every day. We divided the 62 subjects into 3 groups: short-stay (N = 34, residence from 1 to 10 years), middle-stay (N = 18, 11 to 20 years) and long-stay (N = 10, more than 21 years). The mean residence time is 10.1 years (range 1-40).

⁸ We are aware that the population of the two Spanish-speaking groups speaks different diatopic varieties of Spanish. Nevertheless, the *ir a* + infinitive periphrasis is substantially equally used in the Spanish-speaking world, with the exception of the Catalan-speaking regions of the Iberian Peninsula (BLAS ARROYO, 2008).

5. Results

5.1. Italian speakers

The analysis of native speakers' answers allows us to sketch an acceptability gradient for the 24 target sentences, to be used as a baseline. In Table 1, target sentences are ranked from the least to the most acceptable. Figure 1 graphically displays the gradient.

SENTENCE RANKING	SENTENCE	VALUE	MEAN ACCEPTABILITY RATING*
1	<i>Se vai ad andare alla festa, dimmelo!</i>	Planned future	1.01
2	<i>Non trovo le chiavi della macchina: va ad averle prese Luca</i>	Modal	1.03
3	<i>Sono stanco e questa sera non vado a uscire</i>	Planned future	1.05
4	<i>Marco e Ana questa sera vanno a venire al concerto</i>	Future	1.06
5	<i>Se non studi, non vai a passare l'esame</i>	Future	1.19
6	<i>Quel ragazzo va a essere un bravo architetto</i>	Future	1.29
7	<i>Vanno a essere le tre, sono le 14.55</i>	Aspectual (immediate prospective)	1.44
8	<i>È tardi, la biblioteca va a chiudere</i>	Aspectual (immediate prospective)	1.48
9	<i>Senza i guanti vai ad aver freddo alle mani</i>	Future	1.51
10	<i>Vai a credere che sono pazzo, ma non importa!</i>	Modal	1.59

Table 1. *Italian speakers' acceptability judgements*
(* 1 = fully incorrect; 5 = fully correct) (continues).

SENTENCE RANKING	SENTENCE	VALUE	MEAN ACCEPTABILITY RATING*
11	<i>Torniamo a casa, va a piovere</i>	Aspectual (immediate prospective)	1.74
12	<i>Signore e signori, lo spettacolo va a cominciare</i>	Aspectual (immediate prospective)	2.34
13	<i>Vado a raccontarvi la mia giornata tipica in Italia</i>	Aspectual (prospective)	2.57
14	<i>La pasta è pronta e ora vado a mettere il parmigiano</i>	Aspectual (prospective)	3.12
15	<i>Piano piano va a diminuire il numero di contagi</i>	Pleonastic	3.24
16	<i>L'alto numero di morti va a indicare un peggioramento della situazione</i>	Pleonastic	3.37
17	<i>Le voci del coro vanno a formare una bella melodia</i>	Aspectual (culminative)	3.7
18	<i>Questa lezione va a concludere il corso di geografia del professor Rossi</i>	Pleonastic	3.73
19	<i>In questo capitolo vado ad analizzare i dati dello studio</i>	Aspectual (prospective)	3.75
20	<i>Il nuovo palazzo va a sostituire quello vecchio</i>	Aspectual (culminative)	3.8
21	<i>Il denaro va a finanziare il nuovo progetto urbanistico</i>	Aspectual (culminative)	3.85
22	<i>Se non prendi l'ombrello va a finire che ti bagni</i>	Collocation	4.09
23	<i>Vai a sapere cosa pensano gli spagnoli degli italiani</i>	Collocation	4.24
24	<i>Vai a capire che problema ha Lucia: è sempre triste</i>	Collocation	4.3

Table 1. *Italian speakers' acceptability judgements*
 (* 1 = fully incorrect; 5 = fully correct).

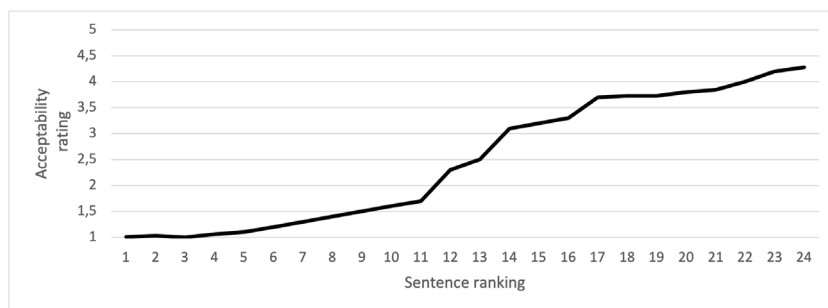


Figure 1. *Italian speakers' acceptability judgement gradient.*

Native speakers' ratings show a clear judgement of unacceptability for those items that are calques of Spanish sentences with (simple and planned) temporal values; the same can be said for the (probability) modal value and for the immediate prospective aspectual value. A t-Test for paired samples run between each sentence reveals a non-significant statistical difference (p always $< .05$) between the judgements for sentence 1 to 11. A first discontinuity in the judgements can be appreciated between sentences 11 and 12: here, the t-Test gives a statistically significant difference (t : -5; df : 163; p : $< .001$). Even though we classify sentence 12 (*Signore e signori, lo spettacolo va a cominciare*) as aspectual (immediate prospective), and even though all the other immediate prospective sentences have been judged as clearly unacceptable, it is from this sentence onwards that the judgements of the Italian native speakers start to rise towards a higher acceptability cline. Sentence 12, which lies, for its semantics, between the immediate prospective and prospective aspectual values, leads on to a cluster of sentences (between 13 to 16) that received doubtful judgments. The mean rating of these sentences is 3.05, with no statistically significant acceptability difference between them: this places pleonastic and aspectual prospective values (displayed by sentences 13 to 16) in a 'grey zone' where the Italian speakers are uncertain about their acceptability. Another discontinuity point runs between sentences 16 and 17: the t-Test reveals a statistically significant difference between the means of the two

sentences (t : -3.7; df : 163; p : < .001). Indeed, from item 17 onwards, Italian native speakers seem to be less unsure about the acceptability of sentences: all aspectual culminative sentences, a pleonastic and an aspectual prospective one were judged above 3.5 points. Only the three collocations were judged above 4 points, showing substantial agreement about their acceptability.

This analysis gives a picture of strong agreement on the unacceptability of temporal, planned future, modal and aspectual immediate prospective values mapped onto *andare a* + infinitive. These values are core values of *ir a* + infinitive, and L1-Spanish speakers of Italian therefore have to unlearn their mapping onto *andare a* + infinitive. Pleonastic and aspectual prospective values are judged with doubt, whereas the aspectual culminative values are accepted with less hesitancy, although they are still rated below 4 points. L1-Spanish speakers therefore need to learn these values as they are not displayed by *ir a* + infinitive but, given native speakers' doubt about their acceptability (which probably reflects their marginality and variability of use), we can infer that the Italian input itself will not give clear cues about their status to the learners. This should not be the case with collocations, which are judged as acceptable with stronger conviction by the Italian speakers. Finally, the prospective value is shared by the two languages, and again, this value should arguably not represent a problem for L1-Spanish speakers, even though the average rating by Italian speakers of aspectual prospective sentences is only 3.2 points. In the next sections we will test these hypotheses.

5.2. *Instructed L1-Spanish speakers*

We compare the judgements of the three competence groups (beginner, intermediate, and advanced) against those of the Italian native speakers.

In Table 2 we report the descriptive statistics, and Figure 2 shows the judgement gradient of the 4 groups.

		BEGINNERS N = 34	INTERMEDIATE N = 31	ADVANCED N = 21	CONTROL (ITALIANS) N = 165
1 - Planned future	Mean	4.02	3.42	1.95	1.01
	Std. Deviation	1.17	1.65	1.28	0.17
2 - Modal	Mean	3.46	2.58	1.76	1.03
	Std. Deviation	1.05	1.41	1.34	0.17
3 - Planned future	Mean	3.91	3.52	1.43	1.05
	Std. Deviation	1.03	1.57	1.08	0.22
4 - Future	Mean	4.11	2.65	1.62	1.06
	Std. Deviation	1.29	1.64	1.2	0.45
5 - Future	Mean	4.12	3.58	1.62	1.19
	Std. Deviation	1.15	1.59	1.16	0.55
6 - Future	Mean	3.78	2.84	1.67	1.29
	Std. Deviation	1.27	1.55	1.24	0.74
7 - Aspectual (immediate prospective)	Mean	3.82	3.16	2.05	1.44
	Std. Deviation	1.17	1.61	1.66	0.84
8 - Aspectual (immediate prospective)	Mean	4.18	3.32	1.95	1.48
	Std. Deviation	1.06	1.62	1.28	0.85
9 - Future	Mean	3.78	2.61	1.81	1.51
	Std. Deviation	1.17	1.54	1.29	0.99
10 - Modal	Mean	3.18	2.74	1.57	1.59
	Std. Deviation	1.36	1.48	1.12	1.12
11 - Aspectual (immediate prospective)	Mean	4.12	3	1.52	1.74
	Std. Deviation	1.12	1.61	0.98	1.19
12 - Aspectual (immediate prospective)	Mean	4.15	3.74	2.48	2.34
	Std. Deviation	1.37	1.67	1.78	1.4

Table 2. *Descriptive statistics of instructed L1-Spanish speakers and Italians (continues).*

		BEGINNERS		INTERMEDIATE		ADVANCED		CONTROL (ITALIANS)	
		N = 34		N = 31		N = 21		N = 165	
13 - Aspectual (prospective)	Mean	3.65		3.23		2.14		2.57	
	Std. Deviation	1.52		1.63		1.39		1.54	
14 - Aspectual (prospective)	Mean	3.38		3.9		2.33		3.16	
	Std. Deviation	1.37		1.3		1.46		1.47	
15 - Pleonastic	Mean	3.38		3.26		2.19		3.24	
	Std. Deviation	1.28		1.39		1.33		1.48	
16 - Pleonastic	Mean	3.62		2.9		1.95		3.37	
	Std. Deviation	1.21		1.4		1.2		1.38	
17 - Aspectual (culminative)	Mean	3.12		3.61		3.62		3.7	
	Std. Deviation	1.37		1.54		1.28		1.26	
18 - Pleonastic	Mean	3.32		3.26		2.76		3.74	
	Std. Deviation	1.32		1.5		1.37		1.26	
19 - Aspectual (prospective)	Mean	3.94		3.55		1.95		3.74	
	Std. Deviation	1.35		1.63		1.47		1.35	
20 - Aspectual (culminative)	Mean	3.41		3.26		2.86		3.8	
	Std. Deviation	1.4		1.44		1.8		1.32	
21 - Aspectual (culminative)	Mean	3.71		3.42		2.76		3.85	
	Std. Deviation	1.38		1.41		1.55		1.26	
22 - Collocation	Mean	2.59		2.58		1.38		4.09	
	Std. Deviation	1.35		1.54		0.67		1.07	
23 - Collocation	Mean	3.35		3.61		3		4.21	
	Std. Deviation	1.41		1.41		1.48		1.03	
24 - Collocation	Mean	3.76		3.32		2.9		4.24	
	Std. Deviation	1.33		1.58		1.81		1.16	

Table 2. *Descriptive statistics of instructed L1-Spanish speakers and Italians.*

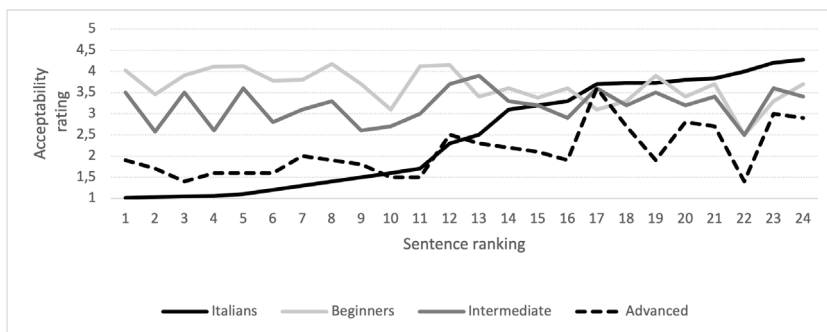


Figure 2. *Italian and instructed L1-Spanish speakers' acceptability judgements.*

We performed a Kruskal-Wallis⁹ test with 'Competence' as the between-subjects independent variable and the judgements of the 24 sentences as dependent variables. When the results of the Kruskal-Wallis test were statistically significant, we performed a Bonferroni-Dunn as post hoc test. Let us consider the first 12 sentences, i.e., those on which are mapped values to be unlearned because they are typical of the Spanish periphrasis, and which the Italian speakers judged as unacceptable. Significant differences (p always $< .05$) for all the 12 items were detected. The post hoc test revealed a significant difference between the judgements of the control group and those of the beginners and intermediate groups for all 12 items: these groups tend to judge these items as acceptable and therefore rate them significantly higher than the Italian native speakers. Between the beginners and intermediate groups few signs of improvement in unlearning are appreciable: indeed, only for sentences 2, 4, 9 and 11 the post hoc revealed a significant difference between the judgements of the two groups. The advanced group is, as we can appreciate from Figure 2, more in line with the trend of the control

⁹ We ran a Kolmogorov-Smirnov test to assess the normal distribution of the 24 dependent variables (judgements about the 24 sentences) across the four levels of the independent variable 'Competence', which includes the three proficiency levels of the Spanish speakers and the Italian informants. The data are not normally distributed (p always $< .05$), and so we opted for non-parametric statistic tests.

group: the judgements of sentences 4, 7, 8, 9, 10, 11 and 12 are not significantly different. Our data suggest that unlearning is at least partially possible, as we witness an accrued ability to recognise the unacceptability of *andare a* + infinitive when associated with immediate prospective and, in some cases, with temporal and modal values. The planned future value seems to be the hardest to unlearn. We note again that this value is the most frequent in Spanish (Aaron, 2006), and this might explain its strongest retention in the advanced learners' interlanguage hypothesis. Any improvement seems very slow and clearly at risk of selective fossilisation: as discussed in Della Putta and Strik-Lievers (2020), improvements happen only at high competence levels.

Moving now to the sentences for which the control group provided the most doubtful ratings (sentences 13 to 16), significant differences between the judgement rate of the four groups are detected: $H(3) = 9.1, p = .04$ for sentence 13; $H(3) = 12.4, p = .01$ for sentence 14; $H(3) = 11.4, p = .02$ for sentence 15; $H(3) = 13.1, p = .01$ for sentence 16. We notice a different phenomenon to that seen for the first 12 sentences: the advanced group judge these items as unacceptable (all 4 judgements are below 2.4 points), whereas the control group raise their judgement rate, albeit remaining essentially in doubt about their acceptability: the post hoc test reveals a significant (p always $< .05$) statistical difference between control and advanced groups for sentences 14, 15 and 16, whereas the judgements for sentence 13 are not significantly different. No significant differences were found between the intermediate and control groups for sentences 14 and 15, neither are they found for sentence 16 between the control and beginners group. Advanced learners seem to judge negatively these more unstable sentences, whereas the basically doubtful trend of beginners and intermediate subjects aligns them more with the judgements of Italian native speakers.

Let us now compare the judgements given to the last 8 sentences (17-24), whose values must be learned by L1-Spanish speakers. In particular, the culminative and pleonastic values are not displayed by *ir a* + infinitive, and they are therefore new uses to be learned.

A non-significant ($H(3) = 5.9, p = .116$) difference for sentence 17 was detected: the four groups all judged it as partially acceptable (mean of the 4 judgements: 3.51). For the remaining seven items, significant differences between the 4 groups are seen. The post hoc test shows that the advanced group judged sentences 16, 18, 19, 20, 21, 22, 23, and 24 always with significantly lower scores than the control group: advanced L1-Spanish speakers seem to overgeneralise the idea that *andare a* + infinitive is not part of the Italian grammar. If compared to the control and advanced groups, the beginner and intermediate groups were more doubtful about the acceptability of the last 8 sentences: apart from sentence 18, on which the judgements of the two groups are in line with those of the control group, for sentences 19-24 the judgements were always statistically significantly lower than the control groups and higher than the advanced groups.

To sum up, the data coming from the instructed population of our study show that unlearning the non-target Italian values of *andare a* + infinitive is possible, but it is a slow process and happens only in advanced learners. Interestingly, advanced learners overgeneralise unlearning even to those sentences that were judged as doubtful or acceptable by the Italians. We can hypothesise that this overgeneralisation is at least partly due to the fact that the positive evidence of the use of *andare a* + infinitive is rather uneven and unclear in the input. A similar but much weaker phenomenon occurs for the beginner and intermediate groups: they judge with greater doubt sentences 13-24, remaining in a sort of 'grey area' where L2 target or L2 partially target sentences are judged neither acceptable nor unacceptable.

5.3. *Non-instructed L1-Spanish speakers*

We first compare the judgements of the three groups (short stay, medium stay, and long stay) against the judgements of the Italian native speakers. In Table 3 we report the descriptive statistics, and Figure 3 shows the judgement gradient of the 4 groups.

		SHORT STAY		MEDIUM STAY		LONG STAY		CONTROL (ITALIANS)	
		N = 34		N = 18		N = 10		N = 165	
1 - Planned future	Mean	2.38	1.33	1	1.01				
	Std. Deviation	1.6	0.97	0	0.17				
2 - Modal	Mean	2.5	1.11	1	1.03				
	Std. Deviation	1.52	0.32	0	0.17				
3 - Planned future	Mean	2.62	1.28	1	1.05				
	Std. Deviation	1.67	0.57	0	0.22				
4 - Future	Mean	2.65	1.11	1	1.06				
	Std. Deviation	1.67	0.32	0	0.45				
5 - Future	Mean	2.56	1.39	1	1.19				
	Std. Deviation	1.74	1.04	0	0.55				
6 - Future	Mean	2.85	1.28	1	1.29				
	Std. Deviation	1.73	0.67	0	0.74				
7 - Aspectual (immediate prospective)	Mean	2.85	1.78	1.2	1.44				
	Std. Deviation	1.54	1.22	0.63	0.84				
8 - Aspectual (immediate prospective)	Mean	2.76	1.17	1.4	1.48				
	Std. Deviation	1.56	0.38	1.26	0.85				
9 - Future	Mean	2.74	1.89	1.2	1.51				
	Std. Deviation	1.58	1.37	0.63	0.99				
10 - Modal	Mean	2.41	1.72	1.8	1.59				
	Std. Deviation	1.56	1.27	1.69	1.12				
11 - Aspectual (immediate prospective)	Mean	3.24	2.28	1.8	1.74				
	Std. Deviation	1.71	1.78	1.69	1.19				
12 - Aspectual (immediate prospective)	Mean	2.76	1.83	1.6	2.34				
	Std. Deviation	1.74	1.29	1.35	1.4				

Table 3. *Descriptive statistics of non-instructed L1-Spanish speakers and Italians (continues).*

		SHORT STAY N = 34	MEDIUM STAY N = 18	LONG STAY N = 10	CONTROL (ITALIANS) N = 165
13 - Aspectual (prospective)	Mean Std. Deviation	3.44 1.5	2.22 1.35	2.4 1.51	2.57 1.54
14 - Aspectual (prospective)	Mean Std. Deviation	2.38 1.58	2.89 1.68	1.7 1.34	3.16 1.47
15 - Pleonastic	Mean Std. Deviation	2.79 1.67	2.61 1.75	2.3 1.34	3.24 1.48
16 - Pleonastic	Mean Std. Deviation	3.12 1.63	3.06 1.66	2.7 1.64	3.37 1.38
17 - Aspectual (culminative)	Mean Std. Deviation	3.09 1.6	2.89 1.57	2.2 1.93	3.7 1.26
18 - Pleonastic	Mean Std. Deviation	2.79 1.75	3.5 1.62	3.6 1.84	3.74 1.26
19 - Aspectual (prospective)	Mean Std. Deviation	3.24 1.63	3.5 1.76	4.5 1.27	3.74 1.35
20 - Aspectual (culminative)	Mean Std. Deviation	2.79 1.34	3.61 1.38	1.8 1.23	3.8 1.32
21 - Aspectual (culminative)	Mean Std. Deviation	3.09 1.69	3.89 1.57	3.4 1.99	3.85 1.26
22 - Collocation	Mean Std. Deviation	4.03 1.19	3.89 1.68	3.9 1.37	4.09 1.07
23 - Collocation	Mean Std. Deviation	3.32 1.63	4.28 1.36	4.6 1.26	4.21 1.03
24 - Collocation	Mean Std. Deviation	3.32 1.63	4.28 1.36	4.6 1.26	4.24 1.16

Table 3. *Descriptive statistics of non-instructed L1-Spanish speakers and Italians.*

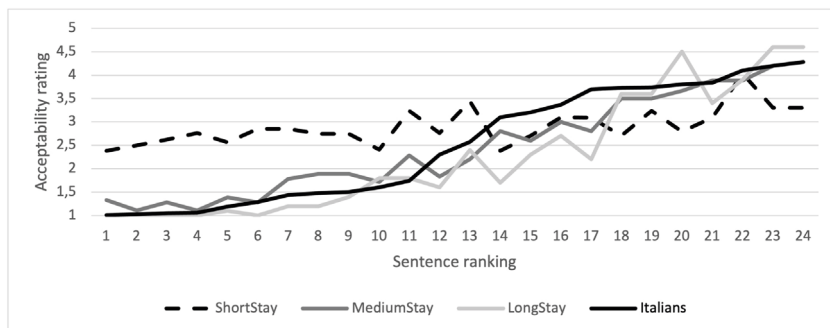


Figure 3. *Italian and non-instructed L1-Spanish speakers' acceptability judgements.*

We performed a Kruskal-Wallis test¹⁰ with 'Length of stay' as the between-subjects independent variable and the judgements of the 24 sentences as dependent variables. When the results of the Kruskal-Wallis test were statistically significant, we performed a Bonferroni-Dunn as post hoc test.

We start our analysis with the first 12 sentences, whose values need to be unlearned. A significant difference between the 4 groups for all the 12 items was seen (p always $< .05$). The post hoc test signalled that between the short-stay and the other three groups the judgement differences are always significant: the short-stay subjects always judged the first 12 items as more acceptable than the other 3 groups. No significant differences were revealed between the control and the long-stay group; the same can be said for the medium stay group, which judged only sentence 9 as significantly more acceptable than the control and long-stay groups. These results suggest that the long-stay group has unlearned the values of *ir a* + infinitive that are not displayed by *andare a* + infinitive. The same can be said for the medium-stay group, who

¹⁰ We ran a Kolmogorov-Smirnov test to assess the normal distribution of the 24 dependent variables (judgements of the 24 sentences) across the four levels of the independent variable 'Length of stay', which includes the three length-of-stay levels of the Spanish speakers and the control group of Italian informants. The data are not normally distributed (p always $< .05$), and so we opted for non-parametric statistic tests.

achieved the same results as the long-stay one, sentence 9 excluded. The short-stay group, on the other hand, always judged the 12 sentences as significantly more acceptable than the control group. Nevertheless, we notice that their judgements exceed 3 (3.24) points only in sentence 11: there is a tendency to judge the first 12 sentences as almost unacceptable. A comparison (performed with a Kruskal-Wallis test with Bonferroni-Dunn as post hoc test) between the short-stay group's judgement of the first 12 items with those of the beginner and intermediate instructed groups revealed that the former significantly (p always < 0.5) outperformed in unlearning the other two. This can be interpreted as an advantage of naturalistic acquisition in comparison to instruction, albeit not targeted to the differences between the two periphrases.

Moving now to the most doubtful sentences for the control group (sentences 13 to 16), significant differences between the 4 groups were detected for all the items: $H(3) = 13.2, p = .01$ for sentence 13; $H(3) = 9.3, p = .04$ for sentence 14; $H(3) = 14.5, p = .002$ for sentence 15; $H(3) = 11.8, p = .03$ for sentence 16, but the post hoc gave a rather complex picture: sentence 13 was judged significantly ($p = .052$) more acceptable by the short-stay group than the other 3 groups; sentence 14 was judged in a significantly more acceptable way by the control group, and the group that best approaches the control judgements is the intermediate one; for sentence 15 and 16 we notice the same phenomenon as for sentence 14: the control group rates them significantly higher than the other 3 groups. The data suggest that the rising gradient of the control group seen in § 5.1 is not followed by the non-instructed speakers: their judgements remain in line with those given to sentences 1 to 12. The exception to this trend is sentence 13, which was judged as more acceptable by the short-stay group. A comparison with the instructed group shows that the advanced and long-stay groups judged very similarly the 4 sentences here: a Mann-Whitney U Test performed on the judgements of the 2 groups detected no significant differences between the two groups: $U = 101, p = .86$ for sentence 13; $U = 85, p = .14$ for sentence 14;

$U = 102$, $p = .88$ for sentence 15; $U = 78$, $p = .21$ for sentence 16. The overgeneralisation of unlearning on these ambiguous sentences happens equally for the two groups. Moreover, the judgements on sentences 14, 15 and 16 by the short and medium-stay groups are always significantly lower than those given by the beginners and intermediate subjects: non-instructed subjects tend to judge the doubtful sentences as less acceptable than the instructed informants. It seems that, apart from sentence 13, rated similarly by all groups, instructed subjects were globally more in line with Italian speakers in judging these items.

For the last cluster of sentences, the values of which need to be learned, no significant differences between the judgements of the 4 groups were revealed for items 17, 20, 22 and 23 (p always $> .05$). These sentences were judged as basically acceptable by all groups, despite the tendency of the short-stay group to give them lower ratings. For sentence 18 and 19, the test signalled a statistical difference ($H(3) = 12.1$, $p = .01$ and $H(3) = 8.2$, $p = .04$, respectively), which the post-hoc reveals to be only between the control and the short-stay group, who rated these items significantly lower than the Italians. For sentence 21 and 24, the significant difference ($H(3) = 18.3$, $p = .003$ and $H(3) = 13.9$, $p = .004$ respectively) was seen between the short-term and the other groups, the short term group being the subjects that judged these items as the least acceptable.

To sum up, our data suggest that the medium- and long-stay groups succeed in both unlearning and learning. In particular, we observe the learning of the target Italian values of *andare* a + infinitive by non-instructed L1-Spanish speakers, while instructed subjects largely failed to learn these values. This suggests that natural acquisition not only helps the unlearning, but also the learning of L2-target values more than instruction, at least when not focused on the phenomenon under consideration. In some cases (sentence 17, 20, 22 and 23) this is appreciable also in short-stay learners, i.e., in people in Italy for 1 to 10 years.

6. *Discussion and conclusions*

Based on the analysis presented in § 5, we can now answer our four research questions:

RQ1 – *What happens when learners deal with a L2 periphrasis that sounds familiar, but whose values and uses are different from those of the ‘corresponding’ L1 one?*

Learners have both to learn new L2-values of the periphrasis and to unlearn L1-like values, by inhibiting the transfer of values that are not target-like. Learning and unlearning are not equally difficult processes for learners (see answer to RQ2), and their success appears to positively correlate with the amount of input exposure (see answer to RQ4); regarding the level of L2-proficiency, the correlation is only positive for unlearning, but not for learning (see answer to RQ3).

RQ2 – *Is the learning of ‘new’ periphrastic functions easier or more difficult than the unlearning of L1-functions that are not target in the L2?*

The unlearning of non-target values of *andare a* + infinitive does not seem to occur in the instructed beginner and intermediate groups: these groups tend to judge as highly acceptable all L1-like values of the periphrasis, therefore showing little awareness of the differences between Spanish and Italian as concerns the structures considered here. For the advanced instructed group, unlearning seems to be easier than learning, to the point, however, that it tends to be overextended: they judge as unacceptable all Spanish-like uses of the periphrasis, including those that are accepted by Italian speakers. Non-instructed subjects show good results in both learning and unlearning: it is especially the long-stay group that demonstrates good developmental trajectories in both tasks. Overall, our study suggests that unlearning is more difficult than learning only for beginner and intermediate instructed subjects.

RQ3 – *Is the ‘attuning’ to target-like Italian features of andare a + infinitive modulated by the proficiency level of the L1-Spanish speakers?*

For the process of unlearning, it is: more advanced learners unlearn much better. For the process of learning, it is not: the more advanced judge *all* uses of the periphrasis as unacceptable, including those accepted by native speakers, whereas the uncertainty of the beginner and intermediate subjects helps them, although probably by chance, to approach more closely Italian native speakers’ ratings.

The answer to this question is therefore nuanced: a higher proficiency level corresponds to judgements that are more native-like regarding features that have to be unlearned, but not for new features that have to be learned.

RQ4 – *Is the ‘attuning’ to target-like Italian features of andare a + infinitive modulated by the amount of input exposure of L1-Spanish speakers?*

For the process of unlearning, we see that after 11 years of exposure (medium- and long-stay groups) there is no difference between the judgements of learners and those of the control group. With less than 11 years exposure (short-stay group), the interlanguage hypotheses are still too L1-biased to approach natives’ judgements. The same can be said for learning, although there are some cases in which the short-stay group already judges the sentences in a way that is not significantly different from native Italians. Overall, we can give a positive answer to this question: the greater the input exposure, the more native-like the judgements.

Some consideration on the limitations of this study are in order. Firstly, our data come from untimed acceptability tests. This implies that the judgements of the sentences may be biased by excessive meta-linguistic introspection and may thus not be fully illustrative of informants’ actual linguistic competence (see Plonsky *et al.*, 2020 for a discussion). Secondly, the test used to create the acceptability gradient of Italian speakers may have induced informants to judge the sentences more strictly, precisely because they were asked to make judgements

on an uncoded and highly diaphasically unstable structure. Production data of Italian speakers might have given a different picture of the use of *andare a* + infinitive. A good way to test if production data differ from acceptability judgements would be to consider, in a future study, corpus data of spoken Italian, in which the values of *andare a* + infinitive could be detected in a bottom-up way; indeed, similar methodological concerns come from Francis (2022: 194-236), who suggests combining judgement data with corpus data.

To sum up, the results of our study are in line with those of scholars who maintain that unlearning is a slow process, as retrieving indirect negative evidence from L2 input is a challenging mental operation that depends on the available amount of L2 input exposure (see discussion in Yang, 2015). This leads to unlearning trajectories that are in line with the process of learning, at least as far as this periphrasis and closely related languages are concerned. With low input exposure, which typically characterises the learning of an L2 in one's home country, unlearning is only reached at an advanced proficiency level and is overgeneralised to target uses of the periphrasis. Such overgeneralisation is due not only to the scarcity of input exposure but also to its quality, which in this case is low given the variability and uneven distribution of *andare a* + infinitive. For these reasons, we suggest that in cases like that investigated here, a targeted pedagogical intervention should be considered to compensate for the lack of input. In other words, structures with similar form and different functions in highly intercomprehensible languages should be presented in a contrastive way in pedagogical grammars and should be explicitly dealt with in the classroom.

Acknowledgements

We would like to thank the colleagues and students who helped us in collecting the data: Maria Vittoria Ambrosini, Mariasole Acquaro and Ana De Hériz. We also thank Gabriele Pallotti, Jacopo Saturno, Alex Housen, and the audience at the international workshop on complexity and difficulty in language acquisition held in Reggio Emilia in November 2021 for their fruitful discussion. Our thanks also go to the two anonymous reviewers for their very constructive comments. Errors and shortcomings are solely our responsibility.

Author contributions

Both authors contributed to the conceptualization of the work. For academic reasons only, the scientific responsibility is attributed as follows: §§ 1 and 2 to Francesca Strik-Lievers, §§ 3 and 5 to Paolo Della Putta, and §§ 4 and 6 to both.

References

- AARON, J. (2006), *Me voy a tener que ir yendo: A corpus-based study of the grammaticization of the ir a+infinitive construction in Spanish*, in SAGARRA, N. and TORIBIO, J. (2006, eds.), *Selected Proceedings of the 9th Hispanic Linguistic Symposium*, Cascadilla Proceedings Project, Somerville, pp. 263-272.
- AMBROSINI, M.V. and DELLA PUTTA, P. (2021), *Problemi di apprendimento di tre perifrasi fra spagnolo e italiano. Osservazioni acquisizionali e proposte pedagogiche*, in «Cuadernos de Filología Italiana», 28, pp. 11-44.
- AMENTA, L. and STRUDSHOLM, E. (2002), «Andare a + infinito» in italiano. *Parametri di variazione sincronici e diacronici*, in «Cuadernos de Filología Italiana», 9, pp. 11-29.
- BAILINI, S. (2016), *La interlingua de lenguas afines. El español de los italianos, el italiano de los españoles*, LED, Milano.
- BERRUTO, G. (2012), *Sociolinguistica dell'italiano contemporaneo*, 2nd ed., Carocci, Roma.
- BERTINETTO, P.M. (1990), *Perifrasi verbali italiane: criteri di classificazione e gerarchia di perifrasticità*, in BERNINI, G. and GIACALONE RAMAT, A. (1990, a cura di), *La temporalità nell'acquisizione di lingue seconde*, FrancoAngeli, Milano, pp. 331-350.
- BERTINETTO, P.M. (1991), *Le perifrasi verbali*, in RENZI, L., SALVI, G. and CARDINALETTI, A. (1991, a cura di), *Grande grammatica italiana di consultazione*, il Mulino, Bologna, pp. 129-161.
- BLAS ARROYO, J.L. (2008), *The variable expression of future tense in peninsular Spanish: The present (and future) of inflectional forms in the Spanish spoken in a bilingual region*, in «Language Variation and Change», 20, pp. 85-126.

- BRAVO, A. (2008), *La perífrasis ir+a+inf en el sistema temporal y aspectual del español*, unpublished doctoral dissertation, Instituto Ortega y Gasset / Universidad Complutense de Madrid.
- BRAVO, A. and LACA, B. (2011), *La perífrasis ir+a+infinitivo y el modo gramatical*, in ESCANDELL, V., LEONETTI, M. and SÁNCHEZ, C. (2011, eds.), *60 problemas de gramática*, Akal, Madrid, pp. 220-226.
- CORDER, P. (1984), *La lingua dell'apprendente*, in ARCAINI, E. and PY, B. (1984, a cura di), *Interlingua, aspetti teorici e implicazioni didattiche*, Istituto della Enciclopedia Italiana, Roma, pp. 49-72.
- CUZA, A., GUIJARRO-FUENTES, P., PIRES, A. and ROTHMAN, J. (2012), *The syntax semantics of bare and definite plural subjects in the L2 Spanish of English natives*, in «International Journal of Bilingualism», 17, 3, pp. 634-652.
- DEKEYSER, R. (2016), *Of moving targets and chameleons: Why the concept of difficulty is so hard to pin down*, in «Studies in Second Language Acquisition», 38, 2, pp. 353-363.
- DELLA PUTTA, P. (2016), *Do we also need to unlearn constructions? The case of constructional negative transfer from Spanish to Italian and its pedagogical implications*, in DE KNOP, S. and GILQUIN, G. (2016, eds.), *Applied Construction Grammar*, De Gruyter, Berlin, pp. 237-270.
- DELLA PUTTA, P. and STRIK-LIEVERS, F. (2020), *From Spanish <estar + gerund> to Italian <stare + gerund>. When teaching to unlearn is needed*, in «Zeitschrift für Interkulturelles Sprachenunterricht», 25, pp. 1407-1430.
- DÍAZ, L., BEL, A. and BEKIOU, K. (2007), *The role of morphological features in the acquisition of Spanish aspectual differences*, in LICERAS, J., ZOBL, H. and GOODLUCK, H. (2007, eds.), *The Role of Formal Features in Second Language Acquisition*, Lawrence Erlbaum, Mahwah, pp. 485-511.
- DIETRICH, W. (1973), *Der periphrastische Verbalaspekt in den romanischen Sprachen*, Niemeyer, Tübingen.
- DIK, S.C. (1997), *The Theory of Functional Grammar: The Structure of the Clause*, De Gruyter, Berlin.
- DONATO, C. and PASQUARELLI-GASCON, V. (2015), *The language of the other: Italian for Spanish speakers through intercomprehension*, in «Itali-ca», 92, 3, pp. 713-735.

- ENRIQUE-ARIAS, A. (2010), *On language contact as an inhibitor of language change. The Spanish of Catalan bilinguals in Majorca*, in BREITBARTH, A., LUCAS, C., WATTS, S. and WILLIS, D. (2010, eds.), *Continuity and Change in Grammar*, John Benjamins, Amsterdam, pp. 97-118.
- FÁBREGAS, A. (2019), *Periphrases in Spanish: Properties, diagnostics and research questions*, in «Borealis. An International Journal of Hispanic Linguistics», 8, 2, pp. 1-82.
- FERNÁNDEZ DE CASTRO, F. (1999), *Las perífrasis verbales en el español actual*, Gredos, Madrid.
- FRANCIS, E. (2022), *Gradient Acceptability and Linguistic Theory*, Oxford University Press, Oxford.
- FROSINI, G. (2020), *Andiamo a... servire la risposta!*, in «Italiano Digitale», 13, 2, pp. 84-86.
- GABRIELE, A. (2009), *Transfer and transition in the SLA of aspect*, in «Studies in Second Language Acquisition», 31, 3, pp. 371-402.
- GARACHANA, M. (2018), *From movement to grammar: Spanish verbal periphrases derived from verbs of movement*, in «Syntaxe et Sémantique», 19, 1, pp. 115-146.
- GASS, S. and MACKAY, A. (2002), *Frequency effects and second language acquisition*, in «Studies in Second Language Acquisition», 24, 2, pp. 249-260.
- GÓMEZ TORREGO, L. (1999), *Los verbos auxiliares. Las perífrasis verbales de infinitivo*, in BOSQUE, I. and DEMONTE, I. (1999, eds.), *Gramática Descriptiva de la Lengua Española*, Espasa Calpe, Madrid, pp. 3323-3390.
- GOOSKENS, C., VAN HEUVEN, J., GOLUBOVIĆ, J., SCHÜPPERT, A., SWARTE, F. and VOIGT, S. (2018), *Mutual intelligibility between closely related languages in Europe*, in «International Journal of Multilingualism», 15, 2, pp. 169-193.
- GRISOLD, T. and KAISER, A. (2017), *Leaving behind what we are not: Applying a system thinking perspective to present unlearning as an enabler for finding the best version of the self*, in «Journal of Organisational Transformation and Social Change», 14, 1, pp. 39-55.
- GRÜTER, T., LIEBERMAN, M. and GUALMINI, A. (2010), *Acquiring the scope of disjunction and negation in L2: A bidirectional study of learners of Japanese and English*, in «Language Acquisition», 17, pp. 127-154.

- HAN, Z. (2014), *From Julie to Wes to Alberto. Revisiting the construct of fossilization*, in HAN, Z. and TARONE, E. (2014, eds.), *Interlanguage. Forty Years Later*, John Benjamins, Amsterdam, pp. 47-74.
- KELLERMAN, E. (1983), *Now you see it, now you don't*, in GASS, S. and SELINKER, L. (1983, eds.), *Language Transfer in Language Learning*, Newbury House, Rowley, pp. 112-134.
- LACA, B. (2021), *Non-passive verbal periphrases in the Romance languages*. *Oxford Research Encyclopedia of Linguistics* [available online at <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-640>, accessed on 12.02.2022].
- LARA-BERMEJO, V. (2016), *La expresión del futuro en las lenguas romances de la Península Ibérica*, in «Boletín de la Real Academia Española», 96, 314, pp. 529-558.
- LARA-BERMEJO, V. (2021), *El futuro morfológico en las lenguas iberorromances. Temporalidad, modalidad y evidencialidad*, in «Revue Romane», 56, 2, pp. 297-326.
- LEFEBVRE, C., WHITE, L. and JORDAN, C. (2006), *Introduction*, in LEFEBVRE, C. WHITE, L. and JORDAN, C. (2006, eds.), *L2 Acquisition and Creole Genesis*, John Benjamins, Amsterdam, pp. 1-14.
- MATTE BON, F. (2006), *Maneras de hablar del futuro en español entre gramática y pragmática. Futuro, “ir a” + infinitivo y presente de indicativo: análisis, usos y valor profundo*, in «Revista Electrónica de Didáctica ELE», 6.
- MCMANUS, K. (2022), *Crosslinguistic Influence and Second Language Learning*, Routledge, New York.
- OLBERTZ, H. (1998), *Verbal Periphrases in a Functional Grammar of Spanish*, De Gruyter, Berlin.
- ORICCO, M. (2020), *Analisi di quattro asimmetrie strutturali fra italiano e spagnolo in alcune grammatiche di italiano per ispanofoni*, unpublished thesis, University of Turin.
- PEARL, L. and MIS, B. (2016), *Induction problems, indirect positive evidence and Universal Grammar: Anaphoric one revisited*, in «Language», 92, pp. 1-30.
- PLONSKY, L., MARSDEN, E., CROWTER, D., GASS, S. and SPINNER, P. (2020), *A methodological synthesis and meta-analysis of judgement tasks in second language research*, in «Second Language Research», 36, 1, pp. 583-621.

- RAE = REAL ACADEMIA ESPAÑOLA (2009), *Nueva gramática de la lengua española*, Espasa Calpe, Madrid.
- RENZI, L. (2012), *Come cambia la lingua. L'italiano in movimento*, il Mulino, Bologna.
- RENZI, L. (2019), *Ancora su come cambia la lingua. Qualche nuova indicazione*, in MORETTI, B., KUNZ, A., NATALE, S. and KRAKENBERGER, E. (2019, a cura di), *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana (Berna, 6-8 settembre 2018)*, Società di Linguistica Italiana, Milano, pp. 13-33.
- RINGBOM, H. (2007), *Cross-linguistic Similarity in Foreign Language Learning*, Multilingual Matters, Clevedon.
- RINGBOM, H. and JARVIS, S. (2009), *The importance of cross-linguistic similarity in foreign language learning*, in LONG, M. and DOUGHTY, C. (2009, eds.), *The Handbook of Language Teaching*, Blackwell, New York, pp. 106-118.
- SCHMID, S. (1994), *L'italiano degli spagnoli. Interlingue di immigrati nella svizzera tedesca*, FrancoAngeli, Milano.
- SCHÜTZE, C. (2019), *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*, Language Science Press, Berlin.
- SCHWARTZ, M. and GOAD, H. (2017), *Indirect positive evidence in the acquisition of a subset grammar*, in «Language Acquisition», 24, pp. 234-264.
- SOBRERO, A. (2003), *Lingue speciali*, in SOBRERO, A. (2003, a cura di), *Introduzione all'italiano contemporaneo. La variazione e gli usi*, Laterza, Bari, pp. 237-277.
- SPADA, N. and TOMITA, Y. (2010), *Interactions between type of instruction and type of language feature: A meta-analysis*, in «Language Learning», 60, 2, pp. 263-308.
- SQUARTINI, M. (1998), *Verbal Periphrases in Romance. Aspect, Actionality and Grammaticalization*, De Gruyter, Berlin.
- STRIK-LIEVERS, F. (2017), *Infinitive con verbi di movimento. Una prima ricognizione fra sincronia e diacronia*, in MAROTTA, G. and STRIK-LIEVERS, F. (2017, a cura di), *Strutture linguistiche e dati empirici in diacronia e sincronia*, Pisa University Press, Pisa, pp. 169-196.

- STRIK-LIEVERS, F. (in press), *Andare a + infinito. Una perifrasi in movimento, fra mutamento e variazione*, in DESSÌ SCHMID, S., FESENMEIER, L. and PACIARONI, T. (in press, Hrsg.), *Movimenti in der italienischen Sprache*, Wissenschaftliche Buchgesellschaft, Darmstadt.
- VALENTINI, A. (2007), *La perifrasi andare a + infinito nell'italiano contemporaneo*, in «Studi Italiani di Linguistica Teorica ed Applicata», 36, 2, pp. 215-234.
- VELAND, R. (2014), *Adjunct/argument alternation without lexical change: Italian a + infinitive attached to a verb of motion*, in «Romanische Forschungen», 126, pp. 153-172.
- YANG, C. (2015), *Negative knowledge from positive evidence*, in «Language», 91, pp. 938-953.
- YUAN, B. (2001), *The status of thematic verbs in the second language-acquisition of Chinese: Against inevitability of thematic-verb raising in second language acquisition*, in «Second Language Research», 17, 2, pp. 248-272.
- ZIRNSTEIN, M., VAN HELL, J. and KROLL, J. (2018), *Cognitive control ability mediates prediction costs in monolinguals and bilinguals*, in «Cognition», 176, pp. 87-106.

PAOLO DELLA PUTTA
Dipartimento di Studi Umanistici
Università di Torino
Via Sant'Ottavio 50
10124 Torino (Italy)
paoloantonio.dellaputta@unito.it

FRANCESCA STRIK-LIEVERS
Dipartimento di Lingue e Culture Moderne
Università di Genova
Piazza Santa Sabina 2
16124 Genova (Italy)
francesca.striklieders@unige.it

Appendix

TARGET ITEMS

1. Se vai ad andare alla festa, dimmelo!
2. Non trovo le chiavi della macchina: va ad averle prese Luca.
3. Sono stanco e questa sera non vado a uscire.
4. Marco e Ana questa sera vanno a venire al concerto.
5. Se non studi, non vai a passare l'esame.
6. Quel ragazzo va a essere un bravo architetto.
7. Vanno a essere le tre, sono le 14.55.
8. È tardi, la biblioteca va a chiudere.
9. Senza i guanti vai ad aver freddo alle mani.
10. Vai a credere che sono pazzo, ma non importa!
11. Torniamo a casa, va a piovere.
12. Signore e signori, lo spettacolo va a cominciare.
13. Vado a raccontarvi la mia giornata tipica in Italia.
14. La pasta è pronta e ora vado a mettere il parmigiano.
15. Piano piano va a diminuire il numero di contagi.
16. L'alto numero di morti va a indicare un peggioramento della situazione.
17. Le voci del coro vanno a formare una bella melodia.
18. Questa lezione va a concludere il corso di geografia del professor Rossi.
19. In questo capitolo vado ad analizzare i dati dello studio.
20. Il nuovo palazzo va a sostituire quello vecchio.
21. Il denaro va a finanziare il nuovo progetto urbanistico.
22. Se non prendi l'ombrello va a finire che ti bagni.
23. Vai a sapere cosa pensano gli spagnoli degli italiani.
24. Vai a capire che problema ha Lucia: è sempre triste.

FILLER ITEMS

1. È stato dormendo fino ad ora ma è ancora stanco.
2. Ho conosciuto a Marta tre anni fa in Inghilterra.
3. Se hai mal di denti devi andare al dentista.
4. Da ore Pedro andava cercando una soluzione al problema.
5. Da quando ha iniziato la dieta è andato costantemente perdendo peso.
6. Non posso venire al cinema, domenica: mio esame è lunedì alle 9.
7. Ho stato in Francia solo una volta nella vita.
8. L'anno scorso Luca ha tornato dalle vacanze molto felice.
9. Da piccolo io era un bambino molto riservato: parlavo poco.
10. Ho stato studiando tutta la notte per l'esame di storia.
11. Tutti i giorni aiuto a mia vicina di casa.
12. L'anno prossimo andrò a Italia per le vacanze.
13. Sabato ho nuotato per due ore senza fermarmi.
14. Lucia ha incontrato Mario per la prima volta due anni fa.
15. Per conoscere il voto dell'esame, sono andato dal professore.
16. Negli ultimi anni, Michele è diventato sempre più critico.
17. Con il tempo, i risultati scolastici di Margherita sono migliorati.
18. La mia macchina è una Fiat Punto del 2005.
19. A 20 anni sono stato in Irlanda per studiare inglese.
20. Io e i miei amici siamo partiti per il Marocco il 10 agosto.
21. Quando andavo a scuola mi piaceva studiare storia.
22. Ieri ho cucinato senza pausa dalle 10 alle 13.
23. Buongiorno, sto cercando Francesca Bianchi, la direttrice dell'ufficio.
24. In Spagna le città sono molto belle e ordinate.

A correlation between applicative marker types and word order patterns: diachronic perspective

DEOKHYUN NAM

ABSTRACT

To date, not enough attention has been paid to applicative marker types in comparison to other parameters yielding the diversity of applicative constructions within and across languages. This study, through analysis of a sample of 50 languages that have applicative markers, uncovered that there is a correlation between applicative marker types and word order patterns. Specifically, applicative prefixes generally co-occur with OV/postpositions whereas applicative suffixes co-occur with both VO/prepositions and OV/postpositions to similar extents. The present study offers a diachronic explanation for this phenomenon by discussing, with examples, the following facts. First, applicative prefixes and applicative suffixes often develop from postpositions and prepositions, respectively. Second, applicative suffixes can develop from second or final verbs (in many cases, “give”) in ‘benefactive applicative periphrases’. Implications of these findings in relation to known correlations concerning word order patterns are also suggested.

KEYWORDS: applicative markers, word order, implicational universals, diachronic typology.

1. *Introduction*

Applicative markers are verbal affixes marking *applicative constructions*, constructions defined as follows:

a means some languages have for structuring clauses which allow the coding of a thematically peripheral argument or adjunct as a core-object argument. Such constructions are signalled by overt verbal morphology. (Peterson, 2007: 1)

Consider the following example from Wolof. Although (1a) and (1b) have the same propositional meanings, they are morphosyntacti-

cally different: “with a spoon” is expressed as an adjunct with the preposition in (1a), while it is expressed as a core-object argument owing to the applicative suffix in (1b). (1b) is an applicative construction. NPs which were made core arguments in applicative constructions, like *kuddu* “spoon” in (1b), are generally called *applied objects* or *applied arguments*.

- (1) Wolof (Atlantic-Congo, Senegal/Gambia/Mauritania)¹

a. *mungi* *lekk* AG *kuddu*
 PRES.3SG eat with spoon
 “He is eating with a spoon.”

(Comrie, 1985: 318)

b. *mungi* *lekk-E* *kuddu*
 PRES.3SG eat-APPL spoon
 “He is eating with a spoon.”

(Comrie, 1985: 318)

Applicative constructions are known to differ from each other across and within languages according to various morphosyntactic or semantic parameters, including: semantic roles of the applied argument, valency of the base verb, obligatoriness of the applicativization, object properties of the applied argument etc. (e.g. Palmer, 1994: 161-168; Payne, 1997: 186-191; Creissels, 2006: 73-84; Peterson, 2007; Kiyosawa and Gerdts, 2010: 329-360; Dixon, 2012: 294-342; Polinsky, 2013; Song, 2018: 383-391; Zúñiga and Kittilä, 2019: 53-70). Of these parameters, the present study focuses on the parameter of *type of applicative marker*. Despite its relative transparency, typological literature has not focused on this parameter as much as one may expect. Correlations of applicative marker types with any grammatical phenomena have never been proposed in previous studies, including Peterson (2007: 172-230) and Polinsky (2013), who deal with a lot of sample languages having applicatives (50 languages and 83 languages, respectively).

¹ Throughout the study, applicative markers and other relevant elements (such as adposition counterparts of the applicative markers) appearing in the examples will be in SMALL CAPS.

The applicative marker used in the Wolof applicative construction (1b) is a suffix. There are also other types of applicative markers attested². One of them is a prefix. An illustration is given below from Winnebago:

- (2) Winnebago (Siouan, Midwestern United States)

a. *kook-EJA naanzhin-je-enan*

box-LOC stand-AUX-DECL

“It is standing in the box.”

(Craig and Hale, 1988: 314, 328)

b. *kook-ra HO-nanzhin-je-enan*

box-DEF INESSIVE-stand-AUX-DECL

“It is standing in the box.”

(Craig and Hale, 1988: 314, 328)

The third type is a circumfix. Nishimoto (2018: 88-90) suggests that Tandroy has an applicative circumfix with the allomorphs of: *añ-* -a, *a-* -añe, *i-* -a and *i-* -añe. The usage is illustrated in (3) below, in which *marare reo* “sick people” is an applied argument that has a beneficiary role:

- (3) Tandroy (Austronesian, Southern Madagascar)

ipaiaÑE vare marare reo

look_for.APPL rice sick_person PL

“(Somebody) looks for rice for sick people.”

(Nishimoto, 2018: 80; transl.: *D.N.*)

The aim of this study is twofold: (i) reveal that applicative marker types correlate with word order patterns by investigating the distribution of different applicative marker types and word order patterns across the world’s languages, and (ii) provide diachronic explanations for the correlation.

The organization of the present paper is as follows: § 2 discusses the language sample; § 3 shows the result of classification for the

² Some applicative markers could be analyzed as clitics (as in Rama; CRAIG and HALE, 1988). However, the present study focuses on the prefix vs. suffix distinction rather than the affix vs. clitic distinction.

sample languages according to the types of applicative markers they have, together with word order patterns in each language; § 4 proposes a correlation deduced from what § 3 shows between applicative marker types and word order (adposition order and constituent order) patterns; § 5 provides some general theoretical foundation for the diachrony of verbal affixes in general; § 6 discusses how the processes whereby adpositions grammaticalize into applicative markers can contribute to explaining the correlation; § 7 discusses how the processes whereby verbs grammaticalize into applicative markers is also compatible with the correlation; § 8 summarizes the diachronic links discussed in §§ 6-7; § 9 briefly discusses how word order phenomena that are different from adposition order may correlate with applicative marker types; § 10 briefly discusses minor diachronic sources of applicative markers for future studies; § 11 provides a conclusion.

2. *Language sample*

I have selected 50 languages with applicative constructions and gathered information about their applicative markers mainly from grammar descriptions to perform the present study. My sample languages are listed below, with information about their genetic affiliations and geographical areas (in a similar fashion to Dryer, 1992)³, together with the major source literature. From a geographic perspective, it can be seen that the sample languages do not show a perfectly equal distribution. Many of the languages are concentrated in the following six specific areas: Sub-Saharan Africa, Island Southeast Asia, Papua New Guinea, Australia, the Americas and the Caucasus, which is somewhat also the case for Peterson's (2007) and Polinsky's (2013) samples of languages with applicatives. From a genealogical perspective, Austronesian languages, Atlantic-Congo languages and Nilotic

³ Although DRYER (1992) integrates North America and Central America into one, I distinguish them. Although genetic affiliations and geographical areas of the sample languages are not relevant to the main argument of the present paper, I introduce them here to show that the sample is not unduly biased in terms of these parameters.

languages occupy a good proportion of the sample. In that way, applicatives themselves are not distributed evenly in the world, but how frequently applicatives are found highly depends on the area and the language family (also, Dixon, 2012: 294 estimates that only «no more than about a quarter of the world's languages» have applicatives). In other words, the concentrated distribution of my sample languages in these particular areas and language families could be seen to some extent as «a scaled-down version» (Song, 2018: 94) of the whole picture of the distribution of applicatives in the world's languages. Note also that the three language families mentioned above are among the biggest ones in the world, contributing to their internal diversity and that languages belonging to different language families in the Americas share some areal features⁴.

⁴ Some of the languages in my sample are not considered to have applicatives by PETERSON (2007) and POLINSKY (2013). These languages will be discussed briefly in what follows. PETERSON (2007) states that Yukaghir does not have applicatives. As I do not know what reasoning there was for this statement (despite his referring to MASLOVA, 1999), I included Kolyma Yukaghir in my sample because it has three applicative suffixes, two discussed by MASLOVA (1999) and NAGASAKI (2003) and the other by MASLOVA (1993: 273). Also, PETERSON (2007) mentions that applicative constructions in the incipient stage of grammaticalization, such as those found in Koyra Chiini (HEATH, 1999), may not be considered prototypical applicatives due to their low productivity. However, I included Koyra Chiini in my sample because what matters is whether or not the element primarily has an applicative function – productivity is not a great issue unless lexicalizations cases are very conspicuous. POLINSKY (2013) classifies, for some reason, Rama as a language without applicatives. However, it is quite safe to consider the Rama prefixes (preverbs such as *ba-* and *yu-*, for example) in question as applicatives, as PETERSON (2007) does, even if CRAIG and HALE (1988) and CRAIG (1990) do not use the term *applicative* itself for these particular elements. Similarly, POLINSKY (2013) classifies Yupik (Central) as a language without applicatives, based on JACOBSON (1995). However, a comprehensive grammar of Central Alaskan Yupik released in a relatively recent year, MIYAOKA (2012), discusses some elements that may be considered applicative affixes found in the language. Therefore, I included Central Alaskan Yupik in my database. Also, PETERSON (2007) and POLINSKY (2013) state that Chechen does not have applicatives. I see the non-core argument-adding verbal prefixes in Chechen-Ingush discussed by NICHOLS (1984; 2011) as applicatives, however, and included the language in my sample. Additionally, preverbs in Indo-European languages are sometimes discussed in terms of their valency-increasing or applicative-like functions, as in Vedic (KULIKOV, 2012), Latin (e.g. LEHMANN, 1983) and German (e.g. ROUSSEAU, 1997: 106-107). However, they were not considered for sampling, because combinations of a preverb and a verb in Indo-European languages tend to show a high degree of lexicalization and because the major typological works on applicatives, including PETERSON (2007), POLINSKY (2013) and ZÚÑIGA and KITILÄ (2019), do not consider Indo-European preverbs at all.

NORTH AMERICA (8 languages): Hualapai (Cochimi-Yuman, Arizona; Ichihashi-Nakayama, 1996), Maricopa (Cochimi-Yuman, Arizona; Gordon, 1986), Winnebago (Siouan, Midwestern United States; Craig and Hale, 1988), Dakota (Siouan, North Dakota/South Dakota; Riggs, 2016 [1852]; Adam, 2019 [1878]), Southern Sierra Miwok (Miwok-Costanoan, California; Freeland, 1951; Broadbent, 1964), Creek (Muskogean, Oklahoma; Martin, 2011), Nez Perce (Sahaptian, Idaho; Rude, 1991), Central Alaskan Yupik (Eskimo-Aleut, Alaska; Miyaoka, 2012)

CENTRAL AMERICA (5 languages): Rama (Chibchan, Nicaragua; Craig and Hale, 1988; Craig, 1990), K'iche' (Mayan, Guatemala; Campbell, 2000), Yucatec Maya (Mayan, Belize/Mexico; Lehmann and Verhoeven, 2006; Lehmann, 2015a), Nahuatl (Uto-Aztecan, Mexico; Andrews, 1975), Southeastern Tepehuan (Uto-Aztecan, Mexico; Willet, 1981)

SOUTH AMERICA (7 languages): Kogi (Chibchan, Colombia; Knuchel, 2020), Kashibo-Kakataibo (Panoan, Peru; Zariquiey Biondi, 2018), Shipibo-Konibo (Panoan, Peru/Brazil; Valenzuela, 2010), Huallaga Quechua (Quechuan, Peru; Weber, 1989), Tariana (Arawakan, Amazonia; Aikhenvald, 2000), Katukina-Kanamari (Harákmbut-Katukinan, Amazonia; Queixalós, 2010; 2014), Mosetén (Mosetén-Chimané, Bolivia; Sakel, 2004)

AFRICA (11 languages): Tandroy (Austronesian, Southern Madagascar; Nishimoto, 2018), Amharic (Afro-Asiatic, Ethiopia; Amberber, 1997; 2000), Koyra Chiini (Songhay, Mali; Heath, 1999), Kipsigis (Nilotic, Kenya; Bii and Lonyangapuo, 2014), Maasai (Nilotic, Kenya/Tanzania; Lamoureux, 2004), Dholuo (Nilotic, Kenya/Tanzania; Okoth-Okombo, 1997; Odera *et al.*, 2017), Wolof (Atlantic-Congo, Gambia/Senegal; Dione, 2013), Mbuun (Atlantic-Congo, Democratic Republic of the Congo [DRC]; Bostoen and Mundeke, 2011), Rwanda (Atlantic-Congo, Rwanda/DRC; Kimenyi, 1980), Swahili (Atlantic-Congo, East Africa; Liu, 2014), Herero (Atlantic-Congo, Namibia/Botswana; Yoneda, 2009)

SOUTHEAST ASIA (6 languages): Rawang (Sino-Tibetan, Myanmar; LaPol-la, 2000), Bantik (Austronesian, Indonesia; Utsumi, 2012), Kambara (Austronesian, Indonesia; Klamer, 1998), Taba (Austronesian, Indonesia; Bowden, 1997), Tukang Besi (Austronesian, Indonesia; Donohue, 1999b; 2001), Javanese (Austronesian, Indonesia; Hemmings, 2013)

AUSTRALIA-NEW GUINEA (7 languages): Warembori (Lower Mamberamo, Indonesia; Donohue, 1999a), Barupu (Skou, Papua New Guinea; Donohue, 2003), Motuna (Bougainville, Papua New Guinea; Onishi, 2000), Kope (Kiwaian, Papua New Guinea; Clifton, 1995), Ngan'gityemmerri (Southern Daly, Australia; Reid, 1990), Warrungu (Pama-Nyungan, Australia; Tsunoda, 1998), Kalkatungu (Pama-Nyungan, Australia; Blake, 1979)

EURASIA (6 languages): Ainu (isolate, Japan; Kindaichi, 1993), Kolyma Yukaghir (Yukaghir, East Siberia; Maslova, 1993; 1999; Nagasaki, 2003), rGyalrong (Sino-Tibetan, Sichuan; Nagano, 2018), Thulung Rai (Sino-Tibetan, Nepal/India; Lahaussois, 2002), Chechen-Ingush (Nakh-Daghestanian, North Caucasus; Nichols, 1984; 2011), Georgian (Kartvelian, Georgia; Creissels, 2006; Kojima, 2012)

3. *Classification*

The sample languages are classified into the four patterns: languages with applicative prefix(es) only (Tables 1 and 2), languages with applicative suffix(es) only (Tables 3 and 4), languages with applicative prefix(es) and suffix(es) (Table 5), and languages with applicative circumfixes. These are shown in turn in the following tables, together with information about applicative marker counts and word order patterns in that language⁵. The word order patterns at issue here include constituent order patterns with regard to V and O (OV order vs. VO order) and adposition order patterns (postpositions vs. prepositions). As demonstrated, in my language sample, OV order overwhelmingly co-occurs with postpositions and VO order with prepositions – a manifestation of well-known implicational universals suggested by Greenberg (1963: 45) and Dryer (1992: 83; 2013; 2019: 65-66) virtually stating that OV order harmonizes with postpositions and VO order with prepositions. The present study will add onto them a correlation between applicative marker types and word order patterns as a new correlation concerning word order patterns.

⁵ The data sources are as indicated in § 2.

LANGUAGE	APPLICATIVE PREFIX NUMBER	BASIC WORD ORDER
Winnebago	4	SOV, postpositions
Dakota	5	SOV, postpositions
Creek	2	SOV, postpositions
Rama	5	SOV, postpositions
Kogi	3	SOV, postpositions
Katukina-Kanamari	4	VSO/SVO, postpositions
Bantik	1	SVO/VOS, prepositions
Kope	1	SOV, postpositions
rGyalrong	1	SOV, postpositions
Ainu	3	SOV, postpositions
Georgian	3	SOV, postpositions
Chechen-Ingush	4	SOV, postpositions
Total: 12/50	Total: 36	

Table 1. *Languages with applicative prefix(es) only.*

APPLICATIVE PREFIX NUMBER	1	2	3	4	5	>5	AVERAGE
LANGUAGE NUMBER	3/12	1/12	3/12	3/12	2/12	0/12	3.0

Table 2. *Distribution of applicative prefix counts.*

LANGUAGE	APPLICATIVE SUFFIX NUMBER	BASIC WORD ORDER
Nez Perce	5	free, postpositions
Southern Sierra Miwok	4	SVO, prepositions
Central Alaskan Yupik	5	SOV, postpositions
Hualapai	1	SOV, postpositions
Southeastern Tepehuan	1	SVO, postpositions
Nahuatl	1	VSO/VOS, relational nouns (postpositions)

Table 3. *Languages with applicative suffix(es) only (continues).*

LANGUAGE	APPLICATIVE SUFFIX NUMBER	BASIC WORD ORDER
K'iche'	1	VOS, prepositions
Yucatec Maya	1	VSO/VOS, prepositions
Tariana	1	free, postpositions
Huallaga Quechua	5	free/SOV, postpositions
Kashibo-Kakataibo	3	SOV, postpositions
Shipibo-Konibo	3	SOV, postpositions
Rwanda	5	SVO, prepositions
Herero	1	SVO, prepositions
Swahili	2	SVO, prepositions
Mbuun	1	SVO, prepositions
Wolof	2	SVO, prepositions
Koyra Chiini	1	SVO, prepositions
Kipsigis	4	VSO/VOS, prepositions
Maasai	3	VSO, prepositions
Dholuo	2	SVO, prepositions
Amharic	2	SOV, postpositions
Rawang	1	SOV(verb-final), postpositions
Barupu	8	SOV, prepositions
Warembori	5	free?, prepositions
Tukang Besi	3	VOS, prepositions
Motuna	1	verb-final, postpositions
Taba	2	SVO, prepositions
Javanese	2	SVO, prepositions
Kambera	1	free, prepositions
Ngan'gityemerri	1	SVO, postpositions
Warrongo	2	free, postpositions
Kalkatungu	1	SOV, postpositions
Thulung Rai	1	SOV, postpositions
Kolyma Yukaghir	3	SOV, postpositions
Total: 35/50	Total: 85	

Table 3. *Languages with applicative suffix(es) only.*

APPLICATIVE SUFFIX NUMBER	1	2	3	4	5	>5	AVERAGE
LANGUAGE NUMBER	15/35	7/35	5/35	2/35	5/35	1/35	about 2.43

Table 4. *Distribution of applicative suffix counts.*

	APPLICATIVE PREFIX NUMBER	APPLICATIVE SUFFIX NUMBER	BASIC WORD ORDER
Maricopa	3	1	SOV, postposition
Mosetén	2	4	SVO, postposition
Total: 2/50	Total: 5	Total: 5	

Table 5. *Languages with both applicative prefix(es) and suffix(es).*

Table 2 and Table 4 can be modified as follows (Tables 6 and 7 respectively), adding in consideration Mosetén and Maricopa:

APPLICATIVE PREFIX NUMBER	1	2	3	4	5	>5	AVERAGE
LANGUAGE NUMBER	3/14	2/14	4/14	3/14	2/14	0/14	about 2.93

Table 6. *Distribution of applicative prefix numbers revisited.*

APPLICATIVE SUFFIX NUMBER	1	2	3	4	5	>5	AVERAGE
LANGUAGE NUMBER	16/37	7/37	5/37	3/37	5/37	1/37	about 2.43

Table 7. *Distribution of applicative suffix numbers revisited.*

The only language in my sample not falling into any of the above groups is Tandroy (1/50), which only has an applicative circumfix with four allomorphs (*añ-* *-a*, *a-* *-añe*, *i-* *-a* and *i-* *-añe*), as already mentioned in § 1.

4. *A correlation between applicative marker types and word order patterns*

The following facts can be observed from the data exhibited in § 3. Of the 14 languages with applicative prefixes, 13 languages (about 92.9%) have predominant OV order or postpositions. There are only 2 languages with predominant VO order or prepositions (Bantik and Mosetén), one of which also has applicative suffixes (Mosetén). In contrast, of the 37 languages with applicative suffixes, 20 languages (about 54.1%) have predominant OV order or postpositions and 19 languages (about 51.4%) have predominant VO order or prepositions. From these observations, the following generalization holds as a correlation between applicative marker types and word order patterns:

(a) *A correlation between applicative marker types and word order patterns*

Languages with applicative prefix(es) are more likely to have postpositions and predominant OV order than languages with applicative suffix(es) are.

More precisely, applicative prefixes generally co-occur with postpositions and predominant OV constituent order whereas applicative suffixes co-occur with both word order patterns to similar extents.

Additionally, there is another relevant observation, concerning the frequency difference between applicative prefixes and applicative suffixes:

(b) *The frequency of applicative prefixes and suffixes*

Applicative suffixes are more frequent than applicative prefixes.

How can the correlation between applicative marker types and word order patterns be explained? To explore this problem, it should be noted that it is not a concomitant effect of any general word order tendency that applies to all grammatical areas (a possibility suggested by an anonymous reviewer). Indeed, it is known that suffixes are more

prevalent than prefixes *on a general basis* (e.g. Bybee *et al.*, 1990) and this seems to be reflected in the classification of the sample languages in the present paper. However, the high frequency of VO order and preposition predominance among languages with applicative prefixes cannot be considered an inevitable outcome of a word order pattern distribution among the world's languages, for the following reasons. Firstly, the proportion of predominantly postpositional (or OV) languages to predominantly prepositional (or VO) languages is roughly equal (e.g. Greenberg, 1963). Secondly, as revealed by Bybee *et al.* (1990), suffixing preference on a general basis is observed across the world's languages without a particularly strong effect on VO languages. Then, what should be questioned is why languages with applicative prefixes exhibit almost uniform OV/postposition predominance, while languages with applicative suffixes are allowed both of OV/postposition predominance and VO/preposition predominance nearly equally, or why VO/preposition predominance is frequently found in conjunction with applicative suffixes but only rarely with applicative prefixes. Another point that supports the assertion that the correlation observed is not a result of a general relationship between adposition order patterns and affix types is suggested by the well-established historical relationship between adposition order patterns and affix types: there are numerous instances in which a case suffix's origin is attributed to a postposition or a case prefix's origin is attributed to a preposition, based on their semantic and phonological similarities (e.g. Heine and Kuteva, 2007: 76, 91-92; Lehmann, 2015b: 84-92). In such instances, there are positive correlations between prefixes and prepositions and between suffixes and postpositions, which contradicts the correlation proposed in the present study with regard to applicative affixes.

Given that situation, it is reasonable to suspect that the type of applicative marker plays a role in the frequent co-occurrence of applicative prefixes and postpositions (or OV) and that of applicative suffixes and prepositions (or VO). If this the case, it means that the types of applicative markers are dependently motivated by their (historical) relationship with adposition order or constituent order. It is possible that there would be independent motivation, in which case the high

frequency of applicative suffixes as mentioned in (b) would not be motivated dependently by any other phenomena but finds an explanation in itself (I owe this suggestion to an anonymous reviewer). However, does the presence of an independent motivation negate the validity of a dependent motivation (or vice versa)? As an example, let us consider suffixing preference on a general basis again. The high frequency of suffixes on a general basis has been given independent explanations, such as a processing explanation (Cutler *et al.*, 1985) or a psycholinguistic explanation (Hawkins and Cutler, 1988). However, Bybee *et al.* (1990) suggest that it also has dependent motivations, specifically that a preposed element is more resistant to morphologization than a postposed element with regard to a stem. The two types of explanations do not negate each other, in the sense that they complementarily explain the same phenomenon from different perspectives. Thus, just as Bybee *et al.* (1990) explain suffixing preference on a general basis with reference to the diachronic developments of the affixes despite the presence of independent motivations, so it is possible to give a diachronic explanation for the high frequency of applicative suffixes over applicative prefixes (this will be done in § 8), irrespective of whether an independent explanation would be possible for it (a possible explanation of such a kind I suppose is, as suggested above already, to see the high frequency of applicative suffixes as an inevitable effect of suffixing preference on a general basis). Such ‘independent explanations’, like a psycholinguistic explanation, can, I suppose, be considered as a kind of synchronic explanation. Another synchronic viewpoint that could be raised against the present study’s claim is an attempt to explain the correlation in synchronic terms. However, the principle mentioned above, that a synchronic explanation (or an independent explanation) and a diachronic explanation can coexist for an identical phenomenon, holds true in this case as well.

Thus, it is appropriate to postulate a diachronic explanation for (a) and (b) based on the historical development of applicative markers. This will be addressed in §§ 6-7, following a discussion of the common theoretical foundation in § 5.

5. *A theoretical foundation for the diachrony of verbal affixes*

The problem of whether an applicative affix is a prefix or suffix is reduced to a problem of morpheme order. The major principle adopted in diachronic studies of the morpheme order problem is that the position of an affix with regard to the verbal root has a strong tendency to reflect the erstwhile position of its source independent word with regard to the verb. In historical approaches to languages, this idea emerged from the observation that affixes develop through syntagmaticization or morphologization of independent words as a result of the frequent placement of these independent words immediately before or after elements of a certain word class. Such ideas were developed and put into practice in a number of diachronic typology studies, including Givón (1971; 2015), Comrie (1980), Bybee (1988), Bybee *et al.* (1990), Siewierska (2010) and Mithun (2011; 2017). In §§ 6-7, it will be discussed how this principle is manifested in the way applicative markers develop from adpositions and verbs, the two most frequent sources of applicative markers (Peterson, 2007: 123-171; Zúñiga and Kittilä, 2019: 222-223). For the issues of less frequent sources, see § 10.

6. *Diachronic links between adpositions and applicative affixes*

Garrett (1990), Baker (1996: 431-432), Creissels (2006: 79), Peterson (2007: 125-129) and Zúñiga and Kittilä (2019: 222-223) discuss that adpositions can be diachronic sources of applicative markers. However, they do not discuss in detail in what mechanisms it is realized or how results are different between postposition source cases and preposition source cases. This will be attempted in the present section.

Consider the following Rama example. According to Craig and Hale (1988), the postposition *ka* as appearing in (4a) is the diachronic source of the applicative prefix *ka-* as appearing in (4b).

(4) Rama (Chibchan, Nicaragua)

- a. *naing taata KA na-ngalbi-u*
 my father PSP/from 1-run-TNS
 “I ran away from my father.”

(Craig, 1990: 127)

- b. *KA-na-ngalbi-u*
 RP/from-1-run-TNS
 “I ran away from him.”

(Craig, 1990: 127, 132)

Likewise, although it is not confirmed in the literature, in Warembori, the applicative suffix *-na* (5b) appears to originate from the preposition *nana* (5a)’s attaching to the preceding verb and reducing its own form.

(5) Warembori (Lower Mamberamo, Indonesia)

- a. *e-na NANA e-me-ro*
 1SG-eat OBL 1SG-house-IND
 “I ate in my house.”

(Donohue, 1999a: 17)

- b. *e-na-NA e-me-ro*
 1SG-eat-APPL 1SG-house-IND
 “I ate in my house.”

(Donohue, 1999a: 17)

Generalization of a similar kind of grammatical change is justified in the following way. Keeping in mind what was discussed in § 5, it should be noted that Dryer (1992: 92-93) discusses a strong correlation between VERB AND ADPOSITIONAL PHRASE order, whereby, in overwhelming cases, OV order co-occurs with PP-V order (see also Baker, 1996: 432) and VO order co-occurs with V-PP order. Also, as already mentioned in the beginning of § 3, in many cases, OV co-occurs with postpositions and VO with prepositions (Greenberg, 1963: 45; Dryer, 1992: 83; 2013; 2019: 65-66). Combining these universals leads to the assumption that, generally, postpositions have many chances to be

placed immediately before the verb and prepositions have many chances to be placed immediately after the verb. Considering finally the fact that the adposition is a closed class and its limited members are used in combination with a wide variety of verbs, it can be postulated that postpositions and prepositions are possible origins of applicative prefixes and applicative suffixes respectively. To sum up, the generalization (c) is gained:

(c) *Adpositional origins of applicative affixes*

The original identity of an applicative prefix is a word that used to frequently come immediately before the verb. Postpositions can fulfill this role in OV (and exceptional VO) languages.

The original identity of an applicative suffix is a word that used to frequently come immediately after the verb. Prepositions can fulfill this role in VO (and exceptional OV) languages.

It should be noted that there are some situations in which a postposition does not come immediately before the verb or a preposition does not come immediately before the verb. Possible situations in which this can happen will be discussed in what follows.

Firstly, if the language in question does not have a rigidly fixed word order, situations will be possible in which a postpositional phrase comes after the verb or a prepositional phrase comes before the verb. However, what is depicted in (c) will be realized to a sufficient degree in such cases as well, as long as the basic or most frequent order allows immediate adjacency of a postposition and the verb or a preposition and the verb and there are many more cases in which a postposition follows the verb or a preposition precedes the verb than the other way around.

Secondly, an adposition heading an adpositional phrase modifying a verb and the verb are not always immediately adjacent even if the word order allows postpositional phrase to follow the verb or prepositional phrase to precede the verb, in that some intervening elements like adverbs, other grammatical arguments or any inflectional morphemes may be present between them. For example, in (6) below, the

NP *šiekar* “sugar” intervenes the preceding postposition *ču* “in” and the verb *tasān* “sprinkle” in the final position:

- (6) Chechen-Ingush (Nakh-Daghestanian, North Caucasus)

čaj-na čU šiekar tasān
tea-DAT in sugar sprinkle
“Put sugar in tea.”

(Nichols, 1984: 193)

In a similar vein, in (7) below, the NP *te bambai* “a comb” is placed between the predicative complex *no-balu* in the initial position and the prepositional phrase headed by the benefactive preposition *ako*.

- (7) *Tukang Besi* (Austronesian, Indonesia)

no-balu te bambai AKO te porai-no
3REAL.S-buy CORE comb BEN CORE fiancée-3GEN
“He bought a comb for his fiancée.”

(Donohue, 2001: 221)

Nevertheless, it is confirmed that these postposition and preposition have developed an applicative suffix and applicative prefix respectively, as in (8) and (9):

- (8) Chechen-Ingush (Nakh-Daghestanian, North Caucasus)

šiekar čaj-na čU-tasān
sugar-NOM tea-DAT in-sprinkle
“Put sugar in tea.”

(Nichols, 1984: 193)

- (9) *Tukang Besi* (Austronesian, Indonesia)

no-balu-AKO te porai-no te bambai
3REAL.S-buy-APPL CORE fiancée-3GEN CORE comb
“He bought a comb for his fiancée.”

(Donohue, 2001: 221)

Why the grammaticalization happens despite such intervention could be explained as follows. If a speaker of a predominantly postpositional language conceives that a postposition heading a

postpositional phrase modifying a verb only has to come before the verb, using the postposition as if it is a prefix to the verb satisfies that demand. In that process, it does not matter if there happens to be any constituent between the postposition and the verb; if the speaker wants to morphologize the postposition to the verb despite the presence of such an intervening constituent, it suffices to place the constituent outside of the resulting prefixed verb. The same will apply to cases of predominantly prepositional languages. Another possibility is, as speculative, that, in an earlier period, the word order rule did not allow it to place a core NP between an adposition and the verb when the applicative marker arose by the adposition's attaching to its adjacent verbs.

Diachronic phenomena that consist with this theoretical background are observable in some more of my sample languages, including both cases in which applicative prefixes likely developed from postpositions and in which applicative suffixes likely developed from prepositions. In §§ 6.1-6.2 below, these will be discussed in turn, as support to the generalization (c). Not every applicative marker has a clear diachronic origin: in some cases, the relationship between the adposition and the applicative affix is not completely evidenced by the source literature but based on my speculations built upon their phonological and semantic similarities. However, as will be seen, there is an abundance of such probable cases, whereas no case is found in which an applicative prefix seems to originate from a preposition or an applicative suffix seems to originate from a postposition. In any case, there will be enough evidence left to justify the general diachronic process of postpositions developing into applicative prefixes and prepositions developing into applicative suffixes.

6.1. *Applicative prefixes from postpositions*

Applicative prefixes originating in postpositions seem to exist in at least 7 languages in my language sample, 4 out of which are from the Americas. These will be illustrated below, in turn.

Riggs (2016 [1852]: 17, 39) and Adam (2019 [1878]: 20, 24) mention that the Dakota applicative prefixes *ki-* (~ *kic'i-*), *a-*, *e-* and

o- probably come from the postpositions *kic'i*, *akan*, *ekta* and *ohna* respectively⁶. Below is illustrated use of the applicative prefix *ki-* ($\sim$ *kic'i-*) (10b, 10c) and the postposition *kic'i* (10a):

(10) Dakota (Siouan, North Dakota/South Dakota)

- a. *he KIC'I mde kta*
him with I.go FUT
“I will go with him.”

(Riggs, 2016 [1852]: 60)

- b. *wowapi KIC'I-caga*
writing for-him.he.made
“He wrote a letter for him.”

(Riggs, 2016 [1852]: 17)

- c. *wowapi KI-caga*
writing to-him.he.made
“He wrote him a letter.”

(Riggs, 2016 [1852]: 17)

I suppose that, in Winnebago, the adverb/postposition *e'gi* “here, in” mentioned in Lipkind (1945: 52), an example cited in (11a) below, might be the origin of the applicative prefix *gi-*, illustrated in (11b). The other Winnebago applicative markers, the prefixes *o-*, *a-* and *i-*, have unknown origins.

(11) Winnebago (Siouan, Midwestern United States)

- a. *hāhē'GI*
hāhē'-E'GI
night-here
“At night.”

(Lipkind, 1945: 52)

⁶ I am not able to judge how plausible these scenarios actually are, with the phonological differences, but I consider at least the relationship between the postposition *kic'i* and the homophonous applicative *kic'i-* quite valid. Also, the remaining applicative marker, the prefix *ic-*, originates from the adverb *ici* “together” (RIGGS, 2016 [1852]: 79; ADAM, 2019 [1878]: 24); it is unknown to me whether it has an adpositional usage.

- b. *cha-a-izhan hin-GI-guch-shannan*
 deer-INDEF IOBJ-DAT-shoot-DECL
 “He shot a deer for me.”

(Craig and Hale, 1988: 331)

According to Gordon (1986: 50), in Maricopa, the case suffixes *-ly*, *-k* and *-m* were the sources from which the semantically and phonologically similar applicative prefixes *ily-*, *k-* and *nym-* emerged, respectively. Note that *-ly*, *-k* and *-m* are described by Gordon as case suffixes rather than postpositions. From this, it is indicated that, for example, there was a common free word source for *-ly* and *ily-*, which could have been a postposition, a relational noun or a serial verb. The same applies to the other two case-applicative pairs. Anyhow, this also matches the generalization depicted in (c). The following are illustrations of *ily-* and *-k* along with their case marker counterparts:

(12) Maricopa (Cochimi-Yuman, Arizona)

- a. *kwnho lames-LY 'shvaw-k*
 basket table-in 1-put-REAL
 “I put the basket on the table.”

(Gordon, 1986: 49)

- b. *mat tdisb ILY-k-shvaw-k*
 earth corn in-IMP-put-REAL
 “Plant the corn in the ground.”

(Gordon, 1986: 50)

- c. *Lynn-sh Yuma-K dii-k*
 Lynn-SBJ Yuma-LOC come-REAL
 “Lynn came from Yuma.”

(Gordon, 1986: 46)

- d. *Heather-sh va-ny-a K-dii-k*
 Heather-SBJ house-DEM-AUGV LOC-come-REAL
 “Heather came from the house.”

(Gordon, 1986: 50)

Craig and Hale (1988) and Craig (1990: 125-126) discuss that all of the Rama applicative prefixes, *ba-*, *yu-*, *ka-*, *su-* and *yaa-*, come from postpositions, which are *ba(ng)*, *u*, *ka(ng)*, *su* and *aa(k)* respectively. Below are illustrations of *u*, *yu*⁷, *su* and *su-*. It seems that a postpositional phrase may either come after the verb or before the verb, as far as I examine Craig and Hale (1988). Note that some degree of lexicalization may be observed in (13b), in that the combination of the verb “come” and the comitative applicative prefix yields the verb “bring”.

(13) Rama (Chibchan, Nicaragua)

- a. *naing taata U n-aakur-u taim ki*
 my father PSP/with I-be-ASP time in
 “I lived with my father at the time.”

(Craig and Hale, 1988: 322)

- b. *nainguku naing taata ngabang*
 thus my father silkgrass
YU-i-siik-i nguu ki
 with-he-come-ASP house in
 “That’s why my father brings the silkgrass in the house.”

(Craig and Hale, 1988: 313)

- c. *nang-SU aa an-kaa-i*
 bed-PSP/on NEG they-put-TNS
 “They don’t put it on the bed.”

(Craig, 1990: 129)

- d. *tauli aa SU-an-kaa-i*
 salt NEG RP/on-they-put-TNS
 “They don’t put salt on it.”

(Craig, 1990: 129)

Of the three Katukina-Kanamari applicative prefixes, *katu-* and *ama-*, come from the postpositions *katu* and *ama* respectively

⁷ CRAIG (1990: 131) mentions that *y-* in *yu-* can be analyzed as a fossilized third person marker or «an epenthetic glide».

(Queixalós, 2010: 41-43; 2014: 303)⁸. Francesc Queixalós (personal communication, 2021) also provides examples of the applicative prefix *to-* and the postposition *-ton*⁹. Although, in each of (14a), (14c) and (14e), the postpositional phrase is placed in the sentence-final position and the postposition is far from the verb, it seems safe to assume that the applicative prefixes in (14b), (14d) and (14f) originate in the respective postpositions, given their phonological and semantic similarities and constituent order flexibility in the language.

(14) Katukina-Kanamari (Harákmbut-Katukinan, Amazonia)

- a. *hoki adu no-KATU*
 talk 1SG 2SG-SOC.INST
 “I am talking to you.”

(Queixalós, 2014: 302)

- b. *i-KATU-hoki i:dik*
 1SG-APPL-talk-2SG 2SG
 “I am talking to you.”

(Queixalós, 2014: 302)

- c. *Dyomi na=donman-na Mayon na=AMA*
 Dyoyomi CASE=go.fishing-DIR Mayon CASE=REC
 “Dyoyomi went fishing for Mayon.”

(Queixalós, 2014: 42)

- d. *Dyomi na=AMA-donman-na Mayon*
 Dyoyomi CASE=APPL-go.fishing-DIR Mayon
 “Dyoyomi went fishing for Mayon.”

(Queixalós, 2014: 42)

- e. *hoki i:dik yo-TON*
 talk you me-to
 “You talked to me.”

(Queixalós, p.c., 2021)

⁸ I owe this information to Francesc Queixalós.

⁹ According to QUEIXALÓS (2010: 43), the remaining applicative marker, the prefix *o-*, does not have a postposition cognate and could date back to the pronoun *o* “other”.

- f. *yo-TO-hoki i:dik*
me-APPL-talk you
“I talked to you.”

(Queixalós, p.c., 2021)

All of the Chechen-Ingush applicative prefixes – *chy-* (~ *ču-*), *t’a-* (*t’y-*), *k’al-* and *dehwa-* – evidently have their origins in their homophonous postpositions as Nichols (1984: 193; 2011: 411-414) discusses. Below are illustrations of *ču-* and *k’al-* (15b, d) and their postposition counterparts (15a, c):

(15) Chechen-Ingush (Nakh-Daghestanian, North Caucasus)

- a. *čaj-na ČU šiekar tasan*
tea-DAT in sugar sprinkle
“Put sugar in tea.”

(Nichols, 2011: 193)

- b. *šiekar čaj-na ČU-tasan*
sugar-NOM tea-DAT in-sprinkle
“Put sugar in tea.”

(Nichols, 2011: 193)

- c. *cyskj istuolaa K’AL iiqqar*
cat table.DAT under run.WP
“The cat ran under the table.”

(Nichols, 2011: 411)

- d. *cyskj istuolaa K’AL-iiqqar*
cat table.DAT under-run.WP
“The cat ran under the table.”

(Nichols, 2011: 411)

Although the diachronic origin of the Ainu multiple-meaning applicative prefix *e-*, as appearing in (16b), is not established, a possible scenario I suppose is that it originally had a postposition-like behavior as in (16a) or was a postposition proper. The same is applied to the Ainu other multiple-meaning applicative prefix, *ko-* (compare (16c) and (16d))¹⁰.

¹⁰ The ultimate origin of *ko-* could be the verb *kor* “have” (KINDAICHI, 1993: 273).

(16) Ainu (isolate, Japan)

- a. *kamuiranke-tam shirka tanne teshpa*
 godgiven-sword bent long squid
kane kutbok-E-chiu

and belt_below-at-arrange

“The sword of the god gift is bent sharply into the spear squid and inserted under the obi, and [...]”

(Kannari and Kindaichi, 1959: 153-154; glosses and transl.: *D.N.*)

- b. *tanekuran nei tonoto a-E-kamuinomi*
 tonight that alcohol 1PL-APPL-pray
kor shiran ruwa-tap-an
 PROG EVD DECL-EMP-COP

“We are going to pray with that alcohol tonight.”

(Kannari and Kindaichi, 1964: 139; glosses and transl.: *D.N.*)

- c. *tush-KO-kira*

rope-with-escape

“(He) ran away tied with the rope.”

(Kindaichi, 1993: 273; glosses and transl.: *D.N.*)

- d. *niwen hokokse KO-horipi*
 fierce cry APPL-dance

“He dances with powerful voice.”

(Kannari and Kindaichi, 1963: 136; glosses and transl.: *D.N.*)

6.2. *Applicative suffixes from prepositions*

Mirror phenomena of what we saw above, that is, applicative suffixes developing from prepositions, likely happened in at least 8 languages in my language sample, including 2 Indonesian languages and 5 African languages.

The first case is Warembori. Although their historical relationship is not mentioned in Donohue (1999a), the prepositions *nana*, *ta* and *tana* in (17a), (17c) and (17e) may be assumed to be the origins of the applicative suffixes *-na*, *-ta* and *-tane* in (17b), (17d) and (17f) respectively:

(17) Warembori (Lower Mamberamo, Indonesia)

- a. *e-na* *NANA* *e-me-ro*
 1SG-eat OBL 1SG-house-IND
 “I ate in my house.”

(Donohue, 1999a: 17)

- b. *e-na-NA* *e-me-ro*
 1SG-eat-APPL 1SG-house-IND
 “I ate in my house.”

(Donohue, 1999a: 17)

- c. *ka-ra-pasi* *TA* *bunupune*
 1PL.INCL-go-all ALL village
 “We all went to the village.” not *“We went to all the villages.”

(Donohue, 1999a: 170)

- d. *ka-ra-pasi-TA* *bunupune*
 1PL.INCL-go-all-APPL village
 “We went to all the villages.” not *“We all went to the village.”

(Donohue, 1999a: 170)

- e. *e-mamieke* *da* *TANA* *Patena*
 1SG-daughter go ALL Mantarbori
 “My daughter’s gone to Mantarbori.”

(Donohue, 1999a: 14)

- f. *e-ra-mo-TANE* *Teba*
 1SG-go-hither-APPL Teba
 “I came from Teba.”

(Donohue, 1999a: 26)

As for *Tukang Besi*, Donohue (1999b: 242, 333; 2001) suggests that the preposition *ako* as appearing in (18a) is the source from which the homophonous applicative suffix *-ako* as appearing in (18b) emerged. Also, it seems to me that the conjunction-preposition *kene* as in (18c) and the applicative suffix *-ngkene* as in (18d) could be related in some way.

(18) *Tukang Besi* (Austronesian, Indonesia)

- a. *no-balu* *te* *bambai* *AKO* *te* *porai-no*
 3REAL.S-buy CORE comb BEN CORE fiancée-3GEN
 “He bought a comb for his fiancée.”

(Donohue, 2001: 221)

- b. *no-balu-AKO* *te* *porai-no* *te* *bambai*
 3REAL.S-buy-APPL CORE fiancée-3GEN CORE comb
 “He bought a comb for his fiancée.”

(Donohue, 2001: 221)

- c. *no-wila* *kua* *koranga-no* *KENE* *porai-no*
 3R.S/A-go ALL garden-3GEN and fiancée-3GEN
 “He went to his garden with his fiancée.”

(Donohue, 2001: 221)

- d. *no-wila-NGKENE* *te* *porai-no* *kua* *koranga-no*
 3R.S/A-go-APPL CORE fiancée-3GEN ALL garden-3GEN
 “He went to his garden with his fiancée.”

(Donohue, 2001: 221)

According to Heath (1999: 137), in *Koyra Chiini*, the instrumental-comitative preposition *nda* that appears in (19a) historically got to be suffixed to verbs like in (19b) as a result of «redrawing of word boundaries» (with some lexicalization effect):

(19) *Koyra Chiini* (Songhay, Mali)

- a. *a-a* *ton* *NDA* *allaa* *feeji* *korey*
 3SG.S-IPFV be.full with just sheep white
 “It was full of nothing but white sheep.”

(Heath, 1999: 157)

- b. *ay* *kaa-NDA* *mana* *attee*
 1SG.S come-with 2SG.DAT tea
 “I have brought some tea for you(SG).”

(Heath, 1999: 157)

Next, Amberber (2000: 321-322) and Creissels (2006: 79) note the similarities of the Amharic applicative suffixes *-ll* and *-(i)bb* to the prepositions *la-* and *ba-* respectively (see also Amberber, 1997: 3-4), and Hudson (2018: 615) mentions that «two prepositions *-bb-* ‘at, on’ and *-ll-* ‘to, for’ are suffixes to verb stems». An illustration of the preposition *ba-* (20a) and the applicative *-ibb* (20b) is shown below. It can be seen that the position of the prepositional phrase with regard to the verb is at odds with the supposed historical process, as in Katukina-Kanamari in (14): word order change might have happened after the applicative suffixes developed, so that the prepositional phrase postposed to the verb is a remnant of the time when prepositional phrases were placed after the verb rather than before the verb¹¹. Another possibility is that, for example, the sequence *-ibb-at* in (20b) is a prepositional phrase that is placed at the end of the predicate and *-et* has a resumptive pronoun function with regard to the preceding topic free NP (*lij-u-n* “the boy”)¹². If the former possibility is the case, the Amharic case will serve as an instantiation of development of applicative suffixes out of prepositions due to the frequent adjacency (in some point in the history) of the prepositions and the verbs. In contrast, if the latter possibility is the case, the suffixal status of the applicative affix will not be relevant to word order; however, as I suppose, it could be the prepositional status of the adposition that makes it easier for the adpositional phrase to come after the verb to fulfill a resumptive function, since, as mentioned in § 6 with reference to Dryer (1992: 92-93), prepositional phrases generally more harmonize with postverbal positions than postpositional phrases do.

(20) Amharic (Afro-Asiatic, Ethiopia)

- a. *astemari-wa* *BƏ-lij-u* *sak ʔ-čč*
 teacher-DEF+F at-boy-DEF laugh+PERF-3F
 “The teacher laughed at the boy.”

(Amberber, 2000: 323)

¹¹ It is known that Amharic underwent VO-to-OV and ModN-to-NMod word order changes (I owe this information to an anonymous reviewer; see BACH, 1970 and GIVÓN, 1971), suggesting that the change supposed here is plausible as well.

¹² I am grateful to an anonymous reviewer who pointed this out.

- b. *astemari-wa lɨj-u-N sak'ə-čč-ɨBB-ət*
 teacher-DEF+F boy-DEF-ACC laugh+PERF-3F-APPL-3M.O
 “The teacher laughed at the boy.”

(Amberber, 2000: 323)

The fifth instance comes from Yucatec Maya. According to Lehmann and Verhoeven (2006) and Lehmann (2015a), some theme applicative constructions by its only applicative suffix *-t* (21b) may be paraphrased using the locative preposition *ti'* (21a). This suggests that the preposition *ti'* could be the source of the applicative suffix *-t*.

(21) Yucatec Maya (Mayan, Belize/Mexico)

- a. *táan u ts'íikil TI' u na'*
 PROG SBJ.3SG feel_angry LOC POSS.3SG mother
 “He is annoyed with / is scolding his mother.”

(Lehmann and Verhoeven, 2006: 471)

- b. *táan u ts'íikil-T-ik u na'*
 PROG SBJ.3SG feel_angry-TRR-INCMPL POSS.3SG mother
 “He is annoyed with / is scolding his mother.”

(Lehmann and Verhoeven, 2006: 471)

According to Kimenyi (1980), the Rwanda locative applicative marker has the allomorphs *-ho*, *-mo* and *-yo*. Of these, at least *-mo* (22e) and *-ho* (22b) appear to me to come from the adpositions *mo* (22d) – or *mú* (22c) – and *ho* (22a) respectively¹³. Although *-ho* seems to only have a postposition cognate, it is likely that it originally had a preposition cognate, from which it emerged. In fact, *-mo* seems to have both postposition and preposition cognates. Compare (22c) and (22d).

(22) Rwanda (Atlantic-Congo, Rwanda/DRC)

- a. *umugóre y-oohere-je isóko HO umubooyi*
 woman she-send-ASP market to cook
 “The woman sent the cook to the market.”

(Kimenyi, 1980: 89)

¹³ KIMENYI (1980: 89) mentions that those adpositions are «underlying prepositions» of the respective applicative markers.

- b. *umugóre y-oohere-jé-HO isóko umubooyi*
 woman she-send-ASP-to market cook
 “The woman sent the cook to the market.”
 (Kimenyi, 1980: 89)
- c. *úmwáana y-a-taa-ye igitabo MÚ máazi*
 child he-PST-throw-ASP book in water
 “The child has thrown the book into the water.”
 (Kimenyi, 1980: 89)
- d. *úmwáana y-a-taa-ye áamázi MO igitabo*
 child he-PST-throw-ASP water in book
 “The child has thrown the book into the water.”
 (Kimenyi, 1980: 89)
- e. *úmwáana y-a-taa-yé-MO áamázi igitabo*
 child he-PST-throw-ASP-in water book
 “The child has thrown the book into the water.”
 (Kimenyi, 1980: 89)

According to Stafford (1967: 16), Okoth-Okombo (1997: 50-56) and Odhiambo and Malherbe (2009: 22-24), Dholuo has the prepositions *ni* (~ *ne*) “for” (23b) and *e* “in” (23d). Okoth-Okombo even shows a locational-copula-like usage of *ni* (~ *ne*) (23a), which I consider may be older than its prepositional usage. It at least seems true that, in (23e), *ni* is a copula and *e* is a preposition, as the cited author’s own gloss indicates. I postulate that these are diachronic sources of the applicative suffixes *-n* expressing “on behalf of”, “towards” and “with regard to” (Odero *et al.*, 2017) as illustrated in (23c) and *-e* expressing “a place where an action takes place” (Odero *et al.*, 2017) as illustrated in (23f) respectively:

- (23) Dholuo (Nilotic, Kenya/Tanzania)
- a. *nyathina NI Nairobi*
 child.mine be.PRES (in) Nairobi
 “My child is in Nairobi.”
 (Okoth-Okombo, 1997: 24)

- b. *ng'ato ong'iewo NE nyathi pala*
 someone buy.PF BEN child knife
 “Someone has bought a knife for the child.”
 (Okoth-Okombo, 1997: 54)
- c. *gi-ndik-o-NE-gi_maber*
 3PLS-write-IND-APPL-3PLO_well
 “They are writing well for them.”
 (Odero *et al.*, 2017: 10)
- d. *E tie mesa*
 in foot table
 “at the foot of the table”
 (Odhiambo and Malherbe, 2009: 23)
- e. *rombe NI E pap*
 sheep.PL COP LOC field
 “The sheeps are in the field.”
 (Okoth-Okombo, 1997: 50)
- f. *i-lem-o-E_mos*
 2SG.S-pray-IND-LOC_silently
 “You are praying silently in a place.”
 (Odero *et al.*, 2017: 11)

The last example is Kipsigis. It has four applicative markers, including *-en*, illustrated in (24b), which is built on (24a). Kipsigis’s related dialect Nandi has the applicative suffix *-e*., illustrated in (25b), which, based on Creider (2002: 176-179)¹⁴, seems to overall cover multiple meanings of the Kipsigis suffix *-en* and originates from the multiple-purpose preposition *e:ng* ($\sim e:n$), which is illustrated in (25a). Also, Kipsigis has prepositions that can paraphrase applicative constructions (Maria Kouneli, personal communication, 2021), including the generic preposition *e:n* “at, to, for” (Driemel and Kouneli, 2021: 13), phonologically and semantically similar to the Kipsigis applicative *-en*. Therefore, it seems safe to assume that the diachronic source of *-en* is a preposition.

¹⁴ I owe the information of CREIDER (2002) to Maria Kouneli.

(24) Kipsigis (Kalenjin, Nilotic, Kenya)

a. *ki-a-um*

PST-1SG/NOM-shade(take shelter)

“I shaded/took shelter.”

(Bii *et al.*, 2014: 306)

b. *ki-a-um-EN* *got*

PST-1SG-shade-INST house.DAT

“I shaded (took shelter) in the house.”

(Bii *et al.*, 2014: 306)

(25) Nandi (Kalenjin, Nilotic, Kenya)

a. *mi: inkwe:k* *ce:pú:nkú:t* *E:N* *tábú:t*

be vegetables.NOM pot in attic

“The vegetables are in the pot *in* the attic.”

(Creider, 2002: 179)

b. *um-E:* *kè:t-í:n*

take.shelter-INST tree-that

“Take shelter in that tree!”

(Creider, 2002: 177)

6.3. *Interim summary*

We saw that postpositions’ becoming morphologized to their following verbs to be applicative prefixes or prepositions’ becoming morphologized to their preceding verbs to be applicative suffixes are plausible in different languages¹⁵. They can be seen as manifestations of Baker’s (1988: 229-304) synchronic theory of ‘Preposition Incorporation’ in the domain of the diachrony. Note that the change from an applicative marker to an adposition is unlikely to occur in that, in general, the latter is the less grammaticalized category in terms of morpho-syntactic dependency and desemanticization. Therefore, the relationship is not circular, at least in the sense that adposition-to-applicative

¹⁵ In these cases, the reason why the applied arguments are core is that adpositions that originally accompanied them with no overt peripheral case government left there by attaching to the verb.

changes seem to be much easier to occur than applicative-to-adposition changes.

On the other hand, in my language sample, not every language shows such a clear historical background of its applicative marker(s) and there are many applicative markers whose origins are unknown. However, while, as shown above, there are several languages in which applicative prefixes may come from postpositions or applicative suffixes may come from prepositions, I find no clear cases to the contrary, i.e. applicative prefixes from prepositions or applicative suffixes from postpositions. Also, when an applicative affix does not have any semantically and phonologically similar adposition in that language, it is difficult to know whether it comes from an adposition. In such a case, however, the possibility is that the applicative affix arose from a postposition or preposition by the same mechanism illustrated in §§ 6.1-6.2 after which it was lost or underwent a drastic semantic change so that it cannot be replaced with the corresponding adposition anymore. Consequently, it is possible to say that there are general diachronic links between applicative prefixes and postpositions and between applicative suffixes and prepositions.

In that way, this section provided a diachronic explanation for the distribution of applicative marker types and word order patterns exhibited in § 3. However, it does not explain the whole picture. The correlation (a) does not state that every language with applicative prefix(es) has predominant OV order and postpositions or every language with applicative suffix(es) has predominant VO order and prepositions. In particular, as seen in § 3, there are many languages which have applicative suffix(es) and OV order and postpositions. The reason why this is the case will be discussed from a historical point of view in § 7.

7. *Diachronic link: applicative suffixes from verbs*

To explain how languages with applicative suffix(es) do not favor any particular word order patterns, it is necessary to consider the other major diachronic source of applicative markers than the adposition,

which is the verb (Haspelmath, 1995: 41-42; Baker, 1996: 431; Garrett, 1990; Peterson, 2007: 130-140; Creissels, 2010; Zúñiga and Kittilä, 2019: 222). Particularly, it would be helpful to see how verbs grammaticalize into benefactive applicative markers. The emergence of a benefactive applicative marker from a verb involves an intermediate stage referred to as «periphrastic benefactive constructions» (Peterson, 2007: 134-135) or «benefactive applicative periphrases» (BAP; Creissels, 2010). As outlined by Creissels (2010: 30), «applicative periphrases are biverbal constructions functionally comparable to monoverbal constructions headed by applicative verb forms» and «the two verbs they involve can be designated as *lexical verb* (abbreviated as Vlex) and *verb-operator* (abbreviated as Vop)». It is known that in many instances of benefactive applicative periphrases, the valency-operator verb is a verb “give” (Creissels, 2006: 79; 2010: 33; Peterson, 2007: 229-230). Creissels (2010) gives a three-way formal distinction of benefactive applicative periphrasis based on whether Vlex or Vop bears the marker: «the serializing type», «the marked-Vop type» and «the marked-Vlex type». The following are examples for each of the three classes:

- (26) «The serializing type»¹⁶: Kana (Cross-river Bantu / Niger-Congo, Nigeria)
Nwíikā wēè ɔ́b títú Nè Nútè
 Nwiika PST roast three_leave_yam give Nute
 “Nwiika roasted a three-leave yam for Nute.”
 (Ikoro, 1996: 254, cited in Creissels, 2010: 39)

- (27) «The marked-Vop type»: Efik (Benue-Congo / Niger-Congo, Nigeria)
nám útóm emì Nò mí!
 do work DEM give 1SG
 “Do this work for me!”
 (Welmers, 1973: 369-370, cited in Creissels, 2010: 39)

¹⁶ CREISSELS’S (2010: 37) definition of *serial verb construction* is as follows: «a complex predicate (i.e., a monoclausal construction involving two or more verbs) showing the following two characteristics: a. no linking element is present between the verbs involved in the construction; b. none of the verbs involved in the construction is in a form implying a non-autonomous status».

- (28) «The marked-Vlex type»: Tamil (Dravidian, India/Sri Lanka)

Rājā Kumār-ukkuk katav-ait tirant-u koṭUTT-ān
 Raajaa Kumaar-DAT door-ACC open-CONV give.PAST-S3SGM
 “Raajaa opened the door for Kumaar.”

(Krishnamurti, 2003: 376, cited in Creissels, 2010: 44)

As suggested in Creissels (2010: 63), applicative affixes are supposed to arise from valency-operator verbs by further progress of grammaticalization. Similarly, Peterson (2007: 134-135) estimates that the Japanese periphrastic benefactive construction like the following which I consider corresponds to Creissels’s (2010: 43) «the marked-Vlex type» is undergoing a process of grammaticalization towards a prototypical applicative construction:

- (29) Japanese (Japonic, Japan)

boku=wa Hanako=ni hon=o kat-te YAT-ta
 I=TOP Hanako=DAT book=ACC buy-CONJ give-PAST
 “I bought a book for Hanako’s sake.”

(Shibatani, 1996: 160, cited in Peterson, 2007: 134)

In the context of my language sample, while some languages have applicative suffixes of an uncertain origin, this development pattern is confirmed for some languages, which will be discussed in the following. One such language is Nez Perce: Rude (1991: 186-187) suggests that the benefactive applicative suffix *-a’n*, as exemplified in (30a) below, is historically related to the verb *’eni* “give”¹⁷, which also gave rise to a benefactive postposition, as seen in (30b):

- (30) Nez Perce (Sahaptin-Klamath, Idaho)

a. *walc paa-ny-A’N-ya ’aayato-na*
 knife 3S.3O-make-BEN-PAST woman-O
 “He made the woman a knife.”

(Rude, 1991: 186)

¹⁷ According to RUDE (1991), all the other applicative markers in Nez Perce (suffixes) also have verbal origins (not “give”).

- b. *walas-na paa-ni-ya 'aayato-ʼAYN*
 knife-O 3S.3O-make-PAST woman-BEN
 “He made a knife for the woman.”

(Rude, 1991: 186)

Secondly, as discussed in § 6.2, the applicative suffix *-ni* (~ *-ne*) in Dholuo, although its immediate source seems to be the preposition *ni*, may have its origins in a copula verb. An additional example of the copula usage of *ni* is provided below:

- (31) Dholuo (Nilo-Saharan, Kenya/Tanzania)

nyathi NI gi buk
 child be with book
 “The child has a book.”

(Okoth-Okombo, 1997: 51)

Thirdly, the *Tukang Besi* applicative suffix *-ako* ultimately traces its origins to the verb *ako* “do for” (Donohue, 1999b: 242, 333), although the prepositional stage (32b) intervenes in some manner between the verbal (32a) and the applicative stages (32c). Another *Tukang Besi* applicative suffix *-ngkene* (32f) also exemplifies this development; its diachronic source seems to be the verb *kene* “accompany” (Donohue, 1999b: 187, 188) (32d), which also developed a prepositional usage (32e).

- (32) *Tukang Besi* (Malayo-Polynesian, Indonesia)

a. *no-wila-AKO-ʼe na ina-no kua daoa*
 3R-go-do.for-3OBJ NOM mother-3POSS ALL market
 “They went for their mother to the market.”

(Donohue, 1999b: 201)

b. *no-ʼema te polisi AKO te ina-no*
 3R-answer CORE policeman BEN CORE mother-3POSS
 “He answered the policeman for his mother.”

(Donohue, 1999b: 227)

- c. *no-wila-AKO te ina-no i daoa*
 3R-go-APPL CORE mother-3POSS OBL market
 “She went to the market for her mother.”
 (Donohue, 1999b: 232)

- d. *no-KENE te ina-no*
 3R-accompany CORE mother-3POSS
 “She accompanied her mother.”
 (Donohue, 1999b: 188)

- e. *no-wila KENE ina-no*
 3R-go accompany mother-3POSS
 “She went with her mother.”
 (Donohue, 1999b: 188)

- f. *no-wila-NGKENE te ina-no*
 3R-go-accompany CORE mother-3POSS
 “She went with her mother.”
 (Donohue, 1999b: 201)

The fourth example is from Barupu. As posited by Donohue (2003: 138), the applicative suffix *-ke* (33b) is «plausibly related to» the verb *ke* “sit” (33a):

- (33) Barupu (Skou, Papua New Guinea)
 a. *bio=venavena k-o-KE-i[sic] pita*
 woman=witch R-<3SG.F>-sit down
 “The witch sat down.”
 (Donohue, 2003: 123)
 b. *a k-u-ai-KE-ni*
 rain R-3SG.F-rain-upon-1SG.F
 “It’s raining on me.”
 (Donohue, 2003: 122)

The fifth example is from Nahuatl. Baker (1996: 431) suggests that the Nahuatl applicative suffix *-lia* (~ *-huia*) as appearing in (34) below, comes from a serial verb “give”.

- (34) Nahuatl (Uto-Aztecan, Mexico)

ni-tē-tla-pāqui-LIA

I-INDEF-INDEF-wash-APPL

“I wash things for people.”

(Andrews, 1975: 107)

Lastly, according to Lyle Campbell (personal communication, 2021), the diachronic source of the instrumental applicative suffix *-b’eh* in K’iche’, illustrated in (35) below, is *-b’eh* with the meanings of “road” and “to go”, the latter being a verb-like meaning.

- (35) K’iche’ (Mayan, Guatemala)

če:ʔ

š-ø-in-č’aya-B’E-x

a:w-e:h

wood ASP-3SG.ABS-1SG.ERG-hit-INST-TR 2SG.POSS-GEN

“I used a stick to hit you.”

(Campbell, 2000: 278)

What should be noted here is that, as suggested by each of the above examples, valency-operator verbs are more likely to result in suffixes than prefixes with regard to the lexical verb, so Creissels (2010: 33) mentions that «irrespective of the status of the language in question with respect to constituent order typology, ‘give’ almost always occupies the second position in BAPs» and concludes that «BAPs using verbs other than ‘give’ in valency operator function, or in which ‘give’ occurs in first position, are exceptional» (Creissels, 2010: 63; for further information, see Creissels, 2006: 79). It is suspected that verbs other than “give” are also likely to occur in the second position¹⁸. Anyway, Creissels’s claim is further supported by the tendency which Bybee *et al.* (1990) discuss accompanies every constituent order pattern based on their extended cross-linguistic survey on «suffixing preference» as already mentioned in § 4, which can be summarized as: with regard to a stem, the following element is more likely to morphologize

¹⁸ The only exceptional language I recognize is Creek, in which the dative applicative prefix *im-* is formally identical with a verb “give” and the instrumental applicative prefix *is-* is formally identical with a verb “take” (Jack B. Martin, personal communication, 2021).

to it to be a suffix than a preceding element is likely to morphologize to it to be a prefix. Concerning serial verbs in particular, Bybee *et al.* (1990: 16) also mention that <in some serial constructions, the second verb is the one to grammaticalize, yielding a postposed gram and perhaps eventually a suffix>.

All that has been stated so far in the present section should be considered in conjunction with the fact that the verb “give” is the most frequent verbal source of applicative markers *in general*, as suggested by Baker (1996: 431), Creissles (2006: 81) and Zúñiga and Kittilä (2019: 222). As a result, it is plausible to conclude that when the diachronic source of an applicative affix is a verb, the applicative marker is more likely to be a suffix than a prefix. This is because the applicative affix is supposed to reflect its original position relative to the head verb, in accordance with the theoretical foundation discussed in § 4 and with Givón’s (2015: 25) aphorism <if today’s bound morphemes are yesterday’s lexical words, then today’s morphology is yesterday’s syntax>. The point is that this holds true regardless of the combination of constituent order and adposition order patterns that the language may possess (including OV/postpositions, VO/prepositions, VO/postpositions and OV/prepositions). Speaking of the six languages in my sample that were cited above, Nez Perce and Nahuatl are postpositional while the other four are prepositional. Thus, it is expected that several other applicative suffixes in my sample may prove to be of verbal origins.

The link between applicative suffixes and Creissels’s (2010) valency-operators is considered also reflected in the following observation: A significant proportion of the languages in my sample classified as having applicative suffixes (16/37 cases or 45.9%) are characterized by the presence of only one applicative suffix (let us call such languages *single-applicative languages*), as illustrated in Table 3. This contrasts with the low frequency of single-applicative languages among those with applicative prefixes, which constitute only 21.4% (3 cases) of the 14 languages in my sample with applicative prefixes, as presented in Table 2. I posit that this contrast can be attributed to the previously mentioned fact that verbs “give” are

the most frequent source of applicative suffixes¹⁹. This presupposes that, in many cases, a language likely has only one highly frequent “give” verb, making it difficult for multiple applicative affixes to develop through that grammaticalization pathway. This situation is contrastive to that of adpositions, in that adpositions are more often than not numerous, each with more or less specific meanings that may warrant grammaticalization into applicative markers, which we saw is the case for many languages in §§ 6.1-6.2. This is further confirmed by the following observation of the sample languages: (i) there are 19 languages with applicative suffixes and postpositions. Out of them, as many as 10 languages (about 52.6%) have only one applicative suffix; (ii) there are 18 languages with applicative suffixes and prepositions. Out of them, only 6 languages (about 33.3%) have only one applicative suffix.

This observation suggests that, when prepositions become applicative suffixes, it may tend to result in more than one applicative suffix per language and, when non-prepositions (in many cases, verbs) become applicative suffixes, it may tend to result in only one applicative suffix per language. Maricopa shows a good contrast by itself, which has three applicative prefixes developed from postpositions (Gordon, 1986: 50), which was discussed in § 6.1, and one benefactive applicative suffix (Gordon, 1986: 85-87) which possibly originated from a valency-operator verb²⁰.

¹⁹ The high frequency of benefactive applicatives in single-applicative languages (e.g. Kalkatungu, cf. BLAKE, 1979, and Thulung Rai, cf. LAHAUSOIS, 2002) may be a further support, in that “give” verbs usually grammaticalize to benefactive applicative markers first (BAKER, 1996: 431; PETERSON, 2007: 229-230; CREISSELS, 2010: 34). This suggests that applicative markers whose origins are not explicitly known may have verbal origins.

²⁰ Also, the Warrongo applicative suffixes *-riL⁽¹⁾* and *-riL⁽²⁾* could be integrated into one, which TSUNODA (1998) differentiates based on the fact that the semantic role of the applied argument is comitative when the base is intransitive whereas it is instrument when the base is transitive. If they are integrated, it will further strengthen the idea I am discussing here.

8. *Summary of the diachronic links*

When a language with applicative suffix(es) has predominant OV order and postpositions, it is likely that the applicative suffix(es) come(s) from valency-operator verb(s) because valency-operator verbs tend to become suffixes rather than prefixes and because the adposition possibility is ruled out. While it is possible that that language formerly had preposition(s) which yielded the applicative suffix(es) and subsequently disappeared, this seems to be an uncommon occurrence, in that word order changes are generally considered more radical than morphologization of an independent word to another.

When a language with applicative suffix(es) has predominant VO order and prepositions, it is likely that both the preposition-to-applicative morphological process and the verb-to-applicative morphological process are equally viable possibilities.

Thus, it is plausible that applicative suffixes generally develop from prepositions or valency-operator verbs, whereas applicative prefixes tend to develop from postpositions, rather than from valency-operator verbs. This discrepancy seems to explain the situation in which there are more languages with applicative suffix(es) ($37/50 = 74\%$) than languages with applicative prefixes ($14/50 = 28\%$), as shown in § 3 using my sample languages and stated in (b) in § 4. And it thus explains why a significant number of languages with applicative suffix(es) exhibit OV order and postpositions, whereas nearly all languages with applicative prefix(es) exhibit VO order and prepositions.

9. *Applicative marker types and word order: further correlations*

As well as the correlation between adposition order patterns and applicative marker types, the correlation between constituent order patterns and applicative marker types can be historically explained, in the following way. First of all, as discussed in Dryer (2019: 66),

the correlation between constituent order patterns and adposition order patterns is historically motivated: Specifically, when verbs are grammaticalized into adpositions, they inevitably result in prepositions in VO languages (for example, as already mentioned in § 7, *Tukang Besi* «is basically a VOS language» (Donohue, 1999b: 189) and its preposition-applicative *ako* is considered being originally a verb (Donohue, 1999b: 242, 333) and in postpositions in OV languages. Then, in turn, prepositions are grammaticalized into applicative suffixes and postpositions into applicative prefixes, so constituent order patterns and applicative marker types prove to be connected via adposition order patterns. Of course, this is not relevant to applicative markers developed from verbs in applicative periphrases.

Dryer (2019: 67-69) also discusses a correlation between GENITIVE AND NOUN order and adposition order patterns whereby languages with GenN order are likely to have postpositions, rather than prepositions, and languages with NGen order are likely to have prepositions, rather than postpositions. This is historically motivated as well, in that adpositions may arise from head nouns in genitive constructions, as Dryer illustrates with English:

- (36) English (Germanic, world language)
- a. *in the side of* > NP *inside* NP
 - b. *by the side of* > NP *beside* NP
 - c. *by the cause of* > NP *because of* NP

(Dryer, 2019: 68)

Based on that, it should be the case as well that GENITIVE AND NOUN order has such a historical connection with types of applicative markers that arise from adpositions.

Demonstrating these connections by directly comparing actual examples of applicative marker types on the one hand and constituent order patterns/GENITIVE AND NOUN order patterns on the other is left for future studies.

10. *Other sources of applicative markers*

In addition to adpositions and verbs, it is known that adverbs, nouns and causative markers may be sources of applicative markers (e.g. Peterson, 2007: 131, 133-141). However, such instances are substantially limited compared with adpositional or verbal source instances (as suggested by Peterson, 2007), so they are unlikely to bring any significant effect on the correlation proposed in the present study. Indeed, my sample languages contain only a handful of such instances. Applicative markers in my sample languages which I acknowledge likely arise from adverbs are: the prefixes *ic-* (Dakota; Adam, 2019 [1878]: 24; Riggs, 2016 [1852]: 79), *gi-* (Winnebago; which also has a related postposition; Lipkind, 1945: 52) and *a-* (Georgian; Harris, 2003: 65). Whether adverbs are more subject to grammaticalization into applicative markers in postpositional (or OV) languages than in prepositional (or VO) languages, whether applicative markers from adverbs tend to be prefixes or suffixes or whether it depends on word order patterns in that language should be investigated in future studies. Applicative markers of plausible nominal origins in my sample languages are: the prefixes *o-* (Ainu; Kindaichi, 1993: 270-271; Bugaeva, 2010: 782, 784) and *im-* (Ngan'gityemerri; Reid, 1990) and the suffix *-b'e* (K'iche'; Lyle Campbell, personal communication, 2021). Applicative markers in my sample languages which are considered to have emerged thorough a reanalysis of causative affixes are: the suffixes *-iish* (Rwanda; Kimenyi, 1988) and *-eesh* (Swahili; Song, 1996: 94), which are apparently related. Applicative markers developed from nouns or causative markers need to be studied with regard to more languages as well.

11. *Conclusion*

Applicative prefixes are likely to co-occur with the OV-postposition combination whereas applicative suffixes co-occur with the VO-preposition combination and the OV-postposition combination to similar extents. This is a correlation related to word order patterns

that has never been proposed, and can be explained by the following facts from a diachronic perspective. First, applicative prefixes and applicative suffixes often develop from postpositions and prepositions respectively, through what can be seen as a historical manifestation of Baker's (1988) 'Preposition Incorporation'. Second, applicative suffixes can develop from second-position (final-position) verbs (often, "give") in periphrastic applicative constructions, in which Creissels's (2010) 'valency-operator' verbs are in many cases postposed to lexical verbs. The second fact is compatible with the observation that, of the languages in my sample that are classified as ones with applicative suffix(es), a good proportion has only one applicative suffix.

Of course, adpositions and verbs are never the only possible diachronic sources of applicative markers, but other sources are known to be possible as well, including nouns, adverbs and causative markers (Peterson, 2007: 131, 133-141), and some languages in and out of my sample have not been provided with information about their applicative markers' origins. However, adpositions and verbs are much more common than them as sources of applicative markers (Peterson, 2007: 123-171). Moreover, the historical links between postpositions and applicative prefixes, between prepositions and applicative suffixes and between valency-operator verbs and applicative suffixes are secure, given the theoretical reasonableness and typological evidence I presented. Thus, it seems safe to say that the correlation claimed primarily stems from the way postpositions, prepositions and verbs develop into applicative markers and which applicative marker types these sources primarily result in respectively.

Finally, there is another implication the present study has to offer. Greenberg (1963: 56-57), Hawkins and Gilligan (1988) and Bybee *et al.* (1990) suggest that OV languages are more likely to have a suffix predominance than VO languages. In historical terms, I consider that one contribution to this tendency is the well-observed processes whereby postpositions morphologize to their preceding governed terms and prepositions morphologize to their following governed terms (which was discussed in § 4). However, the processes the present study focused on entail the opposite directions, whereby postpositions morphologize to their following verbs and prepositions morphologize to their preceding

verbs²¹. Thus, it is expected that, if suffixes and prefixes originating in these pathways, namely applicative affixes of adpositional origins, are excluded from the scope, a stronger correlation between affix positions (affix types) and adposition orders will be gained than has been claimed.

Acknowledgments

This paper is a modified version of a part of my doctoral dissertation (yet to be published). I am very grateful to Sonia Cristofaro, Martin Haspelmath, Fernando Zúñiga and two anonymous reviewers for their valuable comments and suggestions for earlier versions of the present study. I am also obliged to Giovanna Marotta for helping me for a long period of time for the publication of the paper. I thank a lot Lyle Campbell, Maria Kouneli, Jack B. Martin and Francesc Queixalós as well, for personally providing me with helpful information for the study. Great thanks to Jessie Wanner-Kawahara, who revised the style of the paper for me as a matter of course. All remaining errors are mine.

Abbreviations

ABS-absolutive, ACC-accusative, ALL-allative, APPL-applicative, ASP-aspect, AUX-auxiliary, AUGV-augment vowel on nouns, BEN-benefactive, CONV-converb, CONJ-conjunction, COP-copula, DAT-dative, DECL-declarative, DEF-definite, DEM-demonstrative, DIR-direction, EMP-emphasis, ERG-ergative, EVD-evidentiality, F-female, FUT-future, GEN-genitive, IMP-imperative, INCL-inclusive, INCMPL-incomplete, IND-indicative, INDEF-indefinitive, INST-instrumental, IPFV-imperfective, LOC-locative, M-masculine, MOD-modifier, N-noun, NEG-negation, NFUT-nonfuture, NOM-nominative, NP-noun phrase, O-object, OBL-oblique, PERF-perfective, PL-plural, POSS-possessive, PRES-present, PROG-progressive, PSP-postposition, PST-past, REAL-realis, REC-recipient, RP-relational preverb, SBJ-subject, SOC-sociative, SG-singular, TNS-tense, TR-transitive, TRR-transitivizer, WP-witnessed past tense, 1-first person, 2-second person, 3-third person.

²¹ As suggested in § 4, this helps to stress the assumption that the positive correlation between prefixes and postpositions and between suffixes and prepositions basically cannot be extended to other phenomena from applicative markers (however, I suspect a similar correlation between construct marker types, cf. CREISSELS, 2009; 2022, and adposition order patterns to which a similar explanation could be applied).

References

- ADAM, L. (2019, [1878]), *De la dérivation verbal spécifique de l'emboîtement et du polysynthétisme: dans la langue dakota*, in ADAM, L. (2019, [1878]), *Études sur Six Langues Américaines: Dakota, Chibcha, Nahuatl, Kechua, Quiché, Maya*, Wentworth Press, Marrickville, pp. 3-25.
- AIKHENVALD, A.Y. (2000), *Transitivity in Tariana*, in DIXON, R.M.W. and AIKHENVALD, A.Y. (2000, eds.), *Changing Valency*, Cambridge University Press, Cambridge, pp. 145-172.
- AMBERBER, M. (1997), *The unergative-unaccusative distinction and the benefactive applicative in Amharic*, in BAILEY, A.C., MOORE, K.E. and MOXLEY, J.L. (1997, eds.), *Proceedings of the Twenty-third Annual Meeting of the Berkeley Linguistics Society: Special Session on Syntax and Semantics in Africa*, Berkeley Linguistic Society, Berkeley, pp. 1-13.
- AMBERBER, M. (2000), *Valency-changing and valency-encoding devices in Amharic*, in DIXON, R.M.W. and AIKHENVALD, A.Y. (2000, eds.), *Changing Valency*, Cambridge University Press, Cambridge, pp. 312-331.
- ANDREWS, J.R. (1975), *Introduction to Classical Nahuatl*, University of Texas Press, Austin.
- BACH, E. (1970), *Is Amharic an SOV language?*, in «Journal of Ethiopian Studies», 8, 1, pp. 9-20.
- BAKER, M.C. (1988), *Incorporation: A Theory of Grammatical Function Changing*, Chicago University Press, Chicago.
- BAKER, M.C. (1996), *The Polysynthesis Parameter*, Oxford University Press, Oxford.
- BII, J.K., LONYANGAPUO, M. and CHELIMO, A.K. (2014), *Verbal extensions in Kipsigis*, in «International Journal of Humanities and Social Science», 4, 9, 1, pp. 302-311.
- BLAKE, B.J. (1979), *A Kalkatungu Grammar* (Pacific Linguistics Series B, 57), Department of Linguistics, Australian National University, Canberra.

- BOSTOEN, K. and MUNDEKE, L. (2011), *The causative/applicative syncretism in Mbuun (Bantu B87, DRC): Semantic split or phonemic merger?*, in «Journal of African Languages and Linguistics», 32, pp. 179-218.
- BOWDEN, J. (1997), *Taba (Makian Dalam): Description of an Austronesian Language from Eastern Indonesia*, PhD thesis, Department of Linguistics and Applied Linguistics, University of Melbourne, Melbourne.
- BROADBENT, S.M. (1964), *The Southern Sierra Miwok Language*, University of California Press, Berkeley.
- BUGAEVA, A. (2010), *Ainu applicatives in typological perspective*, in «Studies in Language», 34, 4, pp. 749-801.
- BYBEE, J.L. (1988), *The diachronic dimension in explanation*, in HAWKINS, J.A. (1988, ed.), *Explaining Language Universals*, Basil Blackwell, Oxford, pp. 350-379.
- BYBEE, J.L., PAGLIUCA, W. and PERKINS, R.D. (1990), *On the asymmetries in the affixation of grammatical material*, in CROFT, W., DENNING, K. and KEMMER, S. (1990, eds.), *Studies in Typology and Diachrony for Joseph H. Greenberg*, John Benjamins, Amsterdam / Philadelphia, pp. 1-42.
- CAMPBELL, L. (2000), *Valency-changing derivations in K'iche'*, in DIXON, R.M.W. and AIKHENVALD, A.Y. (2000, eds.), *Changing Valency*, Cambridge University Press, Cambridge, pp. 236-281.
- CLIFTON, J.M. (1995), *How many transitivizers are in Kope?*, in ABRAHAM, W., GIVÓN, T. and THOMPSON, S.A. (1995, eds.), *Discourse, Grammar and Typology: Papers in honor of John W. M. Verhaar* (Studies in Language Companion Series, 27), John Benjamins, Amsterdam / Philadelphia, pp. 49-63.
- COMRIE, B. (1980), *Morphology and word order reconstruction: Problems and prospects*, in FISIÁK, J. (1980, ed.), *Historical Morphology* (Trends in Linguistics, Studies and Monographs, 17), De Gruyter Mouton, The Hague, pp. 83-96.
- COMRIE, B. (1985), *Causative verb formation and other verb-deriving morphology*, in TIMOTHY, S. (1985, ed.), *Language Typology and Syntactic Description*. Vol. 3: *Grammatical Categories and the Lexicon*, Cambridge University Press, Cambridge, pp. 309-348.

- CRAIG, C.G. (1990), *A Grammar of Rama*, Université de Lyon, Lyon.
- CRAIG, C.G. and HALE, K. (1988), *Relational preverbs in some languages of the Americas: Typological and historical perspectives*, in «Language», 64, pp. 312-344.
- CREIDER, C.A. (2002), *The semantics of participant types in derived verbs in Nandi*, in «Revue Québécoise de Linguistique», 31, 2, pp. 171-190.
- CREISSELS, D. (2006), *Syntaxe Générale: Une Introduction Typologique*. Vol. 2: *La Phrase* (Collection Langues et Syntaxe), Hermes Sciences, Paris.
- CREISSELS, D. (2009), *Construct forms of nouns in African languages*, in AUSTIN, P. K., NATHAN, D., SELLS, P., BOND, O. and CHARETTE, M. (2009, eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory 2*, SOAS, London, pp. 73-82.
- CREISSELS, D. (2010), *Benefactive applicative periphrasis: A typological approach*, in ZÚÑIGA, F. and KITTLÄ, S. (2010, eds.), *Benefactives and Malefactives: Typological Perspectives and Case Studies*, John Benjamins, Amsterdam / Philadelphia, pp. 29-69.
- CREISSELS, D. (2022), *Binominals and construct marking*, in PEPPER, S., MASINI, F. and MATTIOLA, S. (2022, eds.), *Binominal Lexemes in Cross-Linguistic Perspective: Towards a Typology of Complex Lexemes* (Empirical Approaches to Language Typology, 62), De Gruyter Mouton, Berlin, pp. 73-101.
- CUTLER, A., HAWKINS, J.A. and GILLIGAN, G. (1985), *The suffixing preference: A processing explanation*, in «Linguistics», 23, pp. 723-758.
- DIONE, C.B. (2013), *Valency change and complex predicates in Wolof: An LFG account*, in BUTT, M. and KING, T.H. (2013, eds.), *Proceedings of the LFG 13 Conference*, CSLI Publications, Stanford.
- DIXON, R.M.W. (2012), *Basic Linguistic Theory*. Vol. 3: *Further Grammatical Topics*, Oxford University Press, Oxford.
- DONOHUE, M. (1999a), *Sketchy Grammar of Warembori* (Languages of the World/Materials, 341), Lincom Europa, München.
- DONOHUE, M. (1999b), *A Grammar of Tukang Besi* (Mouton Grammar Library, 20), De Gruyter, Berlin / New York.

- DONOHUE, M. (2001), *Coding choices in argument structure: Austronesian applicatives in texts*, in «Studies in Language», 25, 2, pp. 217-254.
- DONOHUE, M. (2003), *Morphological templates, headedness, and applicatives in Barupu*, in «Oceanic Linguistics», 42, 1, pp. 111-143.
- DRIEMEL, I. and KOUNELI, M. (2021), *Verb-y and noun-y complementation in Kipsigis*, in *On the Nouniness of Propositional Arguments* (43. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, February 25th, 2021), conference handout.
- DRYER, M.S. (1992), *The Greenbergian word order correlations*, in «Language», 68, 1, pp. 81-138.
- DRYER, M.S. (2013), *Relationship between the order of object and verb and the order of adposition and noun phrase*, in DRYER, M.S. and HASPELMATH, M. (2013, eds.), *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig [available online at <http://wals.info/chapter/95>, accessed on 02.07.2021].
- DRYER, M.S. (2019), *Grammaticalization accounts of word order correlations*, in SCHMIDTKE-BODE, K., LEVSHINA, N., MICHAELIS, S.M. and SERŽANT, I.A. (2019, eds.), *Explanation in Typology: Diachronic Sources, Functional Motivations and the Nature of the Evidence* (Conceptual Foundations of Language Science), Language Science Press, Berlin, pp. 63-95.
- FREELAND, L.S. (1951), *Language of the Sierra Miwok* (Indiana University Publications in Anthropology and Linguistics, Memoir 6 of the International Journal of American Linguistics), Waverly Press, Baltimore.
- GARRETT, A. (1990), *Applicatives and preposition incorporation*, in DZIWIŃ, K., FARRELL, P.M. and MEIJAS-BIKANDI, E. (1990, eds.), *Grammatical Relations: A Cross-theoretical Perspective*, CSLI Publications, Stanford, pp. 183-198.
- GIVÓN, T. (1971), *Historical syntax and synchronic morphology: An archeologist's field trip*, in *Papers from the 7th Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago, pp. 394-415.
- GIVÓN, T. (2015), *The Diachrony of Grammar*. Vol. 1, John Benjamins, Amsterdam / Philadelphia.

- GORDON, L. (1986), *Maricopa Morphology and Syntax* (University of California Publications in Linguistics, 108), University of California Press, Berkeley.
- GREENBERG, J.H. (1963), *Some universals of grammar with particular reference to the order of meaningful elements*, in GREENBERG, J.H. (1963, ed.), *Universals of Language*, MIT Press, Cambridge (MA), pp. 40-70.
- HARRIS, A.C. (2003), *Preverbs and their origins in Georgian and Udi*, in BOOIJ, G. and VAN MARLE, J. (2003, eds.), *Yearbook of Morphology 2003*, Kluwer Academic Publishers, Dordrecht, pp. 61-78.
- HASPELMATH, M. (1995), *The converb as a cross-linguistically valid category*, in HASPELMATH, M. and KÖNIG, E. (1995, eds.), *Converbs in Cross-linguistic Perspective: Structure and Meaning of Adverbial Verb Forms - Adverbial Participles, Gerunds* (Empirical Approaches to Language Typology, 13), De Gruyter Mouton, Berlin / New York, pp. 1-55.
- HAWKINS, J.A. and CUTLER, A. (1988), *Psycholinguistic factors in morphological asymmetry*, in HAWKINS, J.A. (1988, ed.), *Explaining Language Universals*, Basil Blackwell, Oxford, pp. 280-320.
- HAWKINS, J.A. and GILLIGAN, G. (1988), *Prefixing and suffixing universals in relation to basic word order*, in «Lingua», 74, pp. 219-259.
- HEATH, J. (1999), *A Grammar of Koyra Chiini: The Songhay of Timbuktu* (Mouton Grammar Library, 19), De Gruyter Mouton, Berlin / New York.
- HEINE, B. and KUTEVA, T. (2007), *The Genesis of Grammar: A Reconstruction* (Studies in the Evolution of Language, 9), Oxford University Press, Oxford.
- HEMMINGS, C. (2013), *Causatives and applicatives: The case for polysemy in Javanese*, in «SOAS Working Papers in Linguistics», 16, pp. 167-194.
- HUDSON, G. (2018), *Amharic*, in COMRIE, B. (2018, ed.), *The World's Major Languages* (3rd ed.), Routledge, London, pp. 611-634.
- ICHIHASHI-NAKAYAMA, K. (1996), *The "applicative" in Hualapai: Its functions and meanings*, in «Cognitive Linguistics», 7, 2, pp. 227-239.
- IKORO, S. (1996), *The Kana Language*, Research School CNWS, Leiden University, Leiden.

- JACOBSON, S.A. (1995), *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*, Alaska Native Language Center, Fairbanks.
- KANNARI, M. 金成 まつ (original texts) and KINDAICHI, K. 金田一京助 (Japanese translation and notes) (1959), *Ainu Jojishi Yūkara Shū* アイヌ叙事詩ユークラ集 [Collection of Ainu epic Yukar]. Vol. 1, Sanseidō 三省堂, Tokyo.
- KANNARI, M. 金成 まつ (original texts) and KINDAICHI, K. 金田一京助 (Japanese translation and notes) (1963), *Ainu Jojishi Yūkara Shū* アイヌ叙事詩ユークラ集 [Collection of Ainu epic Yukar]. Vol. 3, Sanseidō 三省堂, Tokyo.
- KANNARI, M. 金成 まつ (original texts) and KINDAICHI, K. 金田一京助 (Japanese translation and notes) (1964), *Ainu Jojishi Yūkara Shū* アイヌ叙事詩ユークラ集 [Collection of Ainu epic Yukar]. Vol. 4, Sanseidō 三省堂, Tokyo.
- KIMENYI, A. (1980), *A Relational Grammar of Kinyarwanda* (University of California Publications in Linguistics, 91), University of California Press, Berkeley.
- KIMENYI, A. (1988), *Passives in Kinyarwanda*, in SHIBATANI, M. (1988, ed.), *Passive and Voice* (Typological Studies in Language, 16), John Benjamins, Amsterdam / Philadelphia, pp. 355-86.
- KINDAICHI, K. 金田一 京助 (1993), *Ainugo 1 アイヌ語 1* [Ainu 1] (Kindaichi Kyōsuke Zenshū 金田一京助全集 [Complete Works of Kyōsuke Kindaichi], 5), Sanseidō 三省堂, Tokyo.
- KIYOSAWA, K. and GERDTS, D.B. (2010), *Salish Applicatives* (Brill's Studies in the Indigenous Languages of the Americas, 1), Brill, Leiden.
- KLAMER, M. (1998), *A Grammar of Kambara* (Mouton Grammar Library, 18), De Gruyter, Berlin.
- KNUCHEL, D. (2020), *Applicatives in Kogi (Chibchan)*, 53rd Annual Meeting of the Societas Linguistica Europaea (presentation slides).
- KOJIMA, Y. (2012), *Version and object marking in Georgian verbs*, in NAKAMURA, W. and KIKUSAWA, R. (2012, eds.), *Objectivization and Subjectivization: A Typology of Voice Systems* (Senri Ethnological Studies, 77), National Museum of Ethnology, Osaka, pp. 221-235.

- KRISHNAMURTI, B. (2003), *The Dravidian Languages* (Cambridge Language Surveys, 11), Cambridge University Press, Cambridge.
- KULIKOV, L. (2012), *Vedic preverbs as markers of valency-changing derivations: Transitivity and objecthood in Indo-European (evidence from Old Indo-Aryan)*, in «Studies in Language», 36, 4, pp. 721-746.
- LAHAUSOIS, A. (2002), *Aspects of the Grammar of Thulung Rai: An Endangered Himalayan Language*, PhD thesis, University of California, Berkeley.
- LAMOUREAUX, S. (2004), *Applicative Constructions in Maasai*, MA thesis, University of Oregon, Eugene.
- LAPOLLA, R.J. (2000), *Valency-changing derivations in Dulong/Rawang*, in DIXON, R.M.W. and AIKHENVALD, A.Y. (2000, eds.), *Changing Valency*, Cambridge University Press, Cambridge, pp. 282-311.
- LEHMANN, C. (1983), *Latin preverbs and cases*, in PINKSTER, H. (1983, ed.), *Latin Linguistics and Linguistic Theory* (Studies in Language Companion Series, 12), John Benjamins, Amsterdam /Philadelphia, pp. 145-166.
- LEHMANN, C. (2015a), *Valency classes in Yucatec Maya*, in COMRIE, B. and MALCHUKOV, A. (2015, eds.), *Valency Classes in the World's Languages. Vol. 2: Case Studies from Austronesia, the Pacific, the Americas, and Theoretical Outlook* (Comparative Handbooks of Linguistics, 1/2), De Gruyter Mouton, Berlin, pp. 1407-1460.
- LEHMANN, C. (2015b), *Thoughts on Grammaticalization. 3rd edition* (Classics in Linguistics, 1), Language Science Press, Berlin.
- LEHMANN, C. and VERHOEVEN, E. (2006), *Extraversive transitivity in Yucatec Maya and the nature of the applicative*, in KULIKOV, L., MALCHUKOV, A. and DE SWART, P. (2006, eds.), *Case, Valency, and Transitivity*, John Benjamins, Amsterdam / Philadelphia, pp. 465-493.
- LIPKIND, W. (1945), *Winnebago Grammar*, King's Crown Press, New York.
- LIU, J. (2014), *The optionality property of Kiswahili applicatives*, in TEDDIMAN, L. (2014, ed.), *Proceedings of the 2014 Annual Conference of the Canadian Linguistic Association*, Canadian Linguistic Association, Ontario.

- MARTIN, J.B. (2011), *A Grammar of Creek (Muskogee)* (Studies in the Anthropology of North American Indians), University of Nebraska Press, Lincoln.
- MASLOVA, E. (1993), *The causative in Yukaghir*, in COMRIE, B. and POLINSKY, M. (1993, eds.), *Causatives and Transitivity*, John Benjamins, Amsterdam / Philadelphia, pp. 271-285.
- MASLOVA, E. (1999), *A Grammar of Kolyma Yukaghir* (Mouton Grammar Library, 27), De Gruyter Mouton, Berlin.
- MITHUN, M. (2011), *Grammaticalization and explanation*, in HEINE, B. and NARROG, H. (2011, eds.), *Oxford Handbook of Grammaticalization*, Oxford University Press, Oxford, pp. 177-192.
- MITHUN, M. (2017), *Functional approaches*, in LEDGEWAY, A. and ROBERTS, I. (2017, eds.), *The Cambridge Handbook of Historical Syntax*, Cambridge University Press, Cambridge, pp. 687-712.
- MIYAOKA, O. (2012), *A Grammar of Central Alaskan Yupik (CAY)* (Mouton Grammar Library, 58), De Gruyter, Berlin / Boston.
- NAGANO, Y. 長野 泰彦 (2018), *Gyarongo Bunpō Kenkyū* 嘉戎語文法研究 [Study of the rGyalrong Grammar], Kyūkoshoin 汲古書院, Tokyo.
- NAGASAKI, I. 長崎 郁 (2003), *Korima yukagīru-go-niokeru dōshika-o henkō-suru setsubiji* コリマ・ユカギール語における動詞価を変更する接尾辞 [Valency-changing suffixes in Kolyma-Yukaghir], in NAKAGAWA, H. 中川 裕 (2003, ed.), *Yūrashia-shogengo-no Dōshi-ron* ユーラシア諸言語の動詞論 [Topics in Verbs of Eurasian Languages]. Vol. 2, Chiba University, Chiba, pp. 10-27.
- NICHOLS, J. (1984), *Direct and oblique objects in Chechen-Ingush and Russian*, in PLANK, F. (1984, ed.), *Objects: Towards a Theory of Grammatical Relations*, Academic Press, London, pp. 183-209.
- NICHOLS, J. (2011), *Ingush Grammar* (University of California Publications in Linguistics, 143), University of California Press, Berkeley.
- NISHIMOTO, N. 西本 希呼 (2018), *Ibara-no Kuni-no Gengo: Madagasukaru-nanbu Tanruigo-no Kijutsu* 茨の国の言語: マダガスカル南部タンルイ語の記述 [A Language of a Briers' Land: A Description of Tandroy of Southern Madagascar], Keio University Press, Tokyo.

- ODERO, E.O., SANDE, W.O., OGONE, O. and ODONDO, E. (2017), *Syntactic properties of Dholuo verbal extensions*, in «Journal of Literature, Languages and Linguistics», 35, pp. 9-15.
- ODHIAMBO, N. and MALHERBE, M. (2009), *Parlons Luo: Langue du Kenya*, Éditions L'Harmattan, Paris.
- OKOTH-OKOMBO, D. (1997), *A Functional Grammar of Dholuo (Nilo-Saharan 12)*, Rüdiger Köppe Verlag, Köln.
- ONISHI, M. (2000), *Transitivity and valency-changing derivations in Motuna*, in DIXON, R.M.W. and AIKHENVALD, A.Y. (2000, eds.), *Changing Valency*, Cambridge University Press, Cambridge, pp. 115-144.
- PALMER, F.R. (1994), *Grammatical Roles and Relations* (Cambridge Textbooks in Linguistics), Cambridge University Press, Cambridge.
- PAYNE, T.E. (1997), *Describing Morphosyntax: A Guide for Field Linguists*, Cambridge University Press, Cambridge.
- PETERSON, D.A. (2007), *Applicative Constructions*, Oxford University Press, Oxford.
- POLINSKY, M. (2013), *Applicative constructions*, in DRYER, M.S. and HASPELMATH, M. (2013, eds.), *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig [available online at <http://wals.info/chapter/109>, accessed on 28.06.2021].
- QUEIXALÓS, F. (2010), *Grammatical relations in Katukina-Kanamari*, in GILDEA, S. and QUEIXALÓS, F. (2010, eds.), *Ergativity in Amazonia* (Typological Studies in Language, 89), John Benjamins, Amsterdam / Philadelphia, pp. 235-284.
- QUEIXALÓS, F. (2014), *Incremental valence changes in an ergative vs. an accusative language: A comparison between Katukina-Kanamari and Sikuani*, in QUEIXALÓS, F., TELLES, S. and BRUNO, A.C. (2014, eds.), *Valence Increase in Amazonian Languages*, Instituto Caro and Cuervo, Bogotá, pp. 295-309.
- REID, N. (1990), *Ngan'gityemerri: A Language of the Daly River Region, Northern Territory of Australia*, PhD thesis, Australian National University, Canberra.

- RIGGS, S.R. (2016 [1852¹], ed.), *Grammar and Dictionary of the Dakota Language: Collected by the Members of the Dakota Mission*, Wentworth Press, Marrickville.
- ROUSSEAU, A. (1997), *La double transitivité existe-t-elle?*, in ROUSSEAU, A. (1997, éd.), *La Transitivité*, Presses Universitaires du Septentrion, Ville-neuve-d'Ascq, pp. 85-112.
- RUDE, N. (1991), *Verbs to promotional suffixes in Sahaptian and Klamath*, in TRAUGOTT, E.C. and HEINE, B. (1991, eds.), *Approaches to Grammaticalization*. Vol. 2: *Focus on Types of Grammatical Markers*, John Benjamins, Amsterdam / Philadelphia, pp. 185-199.
- SAKEL, J. (2004), *A Grammar of Moseṭén* (Mouton Grammar Library, 33), De Gruyter Mouton, Berlin / New York.
- SHIBATANI, M. (1996), *Applicatives and benefactives: A cognitive account*, in SHIBATANI, M. and THOMPSON, S.A. (1996, eds.), *Grammatical Constructions: Their Form and Meaning*, Oxford University Press, Oxford, pp. 157-194.
- SIEWIERSKA, A. (2010), *Person forms*, in SONG, J.J. (2010, ed.), *The Oxford Handbook of Linguistic Typology*, Oxford University Press, Oxford, pp. 322-343.
- SONG, J.J. (1996), *Causatives and Causation*, Routledge, London.
- SONG, J.J. (2018), *Linguistic Typology*, Oxford University Press, Oxford.
- STAFFORD, R.L. (1967), *Elementary Luo Grammar, with Vocabularies*, Oxford University Press, Nairobi.
- TSUNODA, T. (1998), *Applicative constructions in Warrungu of Australia*, in SIEWIERSKA, A. and SONG, J.J. (1998, eds.), *Case, Typology and Grammar: In honor of Barry J. Blake* (Typological Studies in Language, 38), John Benjamins, Amsterdam / Philadelphia, pp. 343-373.
- UTSUMI, A. (2012), *Applicative verbs and applicative construction in the Bantik language*, in «Journal of the Southeast Asian Linguistics Society», 5, pp. 107-125.
- VALENZUELA, P.M. (2010), *Applicative constructions in Shipibo-Konibo (Panoan)*, in «International Journal of American Linguistics», 76, 1, pp. 101-144.

- WEBER, D.J. (1989), *A Grammar of Huallaga (Huánuco) Quechua* (University of California Publications in Linguistics, 112), University of California Press, Berkeley.
- WELMERS, W.E. (1973), *African Language Structures* (University of California Publications in Linguistics), University of California Press, Berkeley.
- WILLETT, T.L. (1981), *Advancements to direct object in Southeastern Tepehuan*, in «Work Papers of the Summer Institute of Linguistics, University of North Dakota Session», 25, 4, pp. 59-74.
- YONEDA, N. 米田 信子 (2009), *Hererogo-niokeru tekiyōkeikōbun-to mokutekigo-no taishōsei* ヘレロ語における適用形構文と目的語の対称性 [Applicative construction and object a/symmetry in Herero (Bantu, R31)], in «Asian and African Languages and Linguistics», 4, pp. 5-37.
- ZARIQUIEY BIONDI, R. (2018), *A Grammar of Kashibo Kakataibo* (Mouton Grammar Library, 75), De Gruyter Mouton, Berlin.
- ZÚÑIGA, F. and KITTLÄ, S. (2019), *Grammatical Voice* (Cambridge Textbooks in Linguistics), Cambridge University Press, Cambridge.

DEOKHYUN NAM
 Dipartimento di Studi Umanistici
 Università di Pavia
 Corso Strada Nuova 65
 27100 Pavia (Italy)
 deokhyun.nam01@universitadipavia.it



Armenian stem classes and the Western-Middle Iranian oblique plural case

MARCO FATTORI

ABSTRACT

So far, no definitive explanation has been given for the fact that Iranian loanwords in Classical Armenian are very often assigned to a morphological class reflecting the corresponding Old Iranian declension. This has rightly been regarded as problematic because the loanwords entered Armenian in the Middle Iranian period, when Western Iranian dialects had already lost final syllables. In this article, it is argued that the correspondence between Armenian and Iranian stem vowels could be inferred from the oblique plural case, which probably preserved traces of the Old Iranian declension still in the 1st century AD.

KEYWORDS: Armenian, Iranian, loanwords, stem classes.

1. Introduction

Since the very first steps of the research on Iranian borrowings in Armenian, scholars noticed that in a considerable number of cases the Armenian declension in which Iranian loanwords are integrated seems to reflect the corresponding O(ld) Ir(anian) stem vowel.

So, for example, Arm. *dew*, *diwac* ‘demon’, *ašxarh*, *°ac* ‘land’, *tačar*, *°ac* ‘palace’ correspond respectively to OIr. **daiva-* (OP *daiva-*, Av. *daēuua-*), **xšaθra-* (Av. *xšaθra-*) and **tacara-* (OP *tacara*); Arm. *axt*, *°ic* ‘sickness’, *baxt*, *°ic* ‘fortune’ and *hamboyr*, *hambowric* ‘kiss’ correspond respectively to OIr. **axti-* (Av. *axti-*), **baxti-* (Ved. *bhakti-*) and **hambaudi-* (Av. *baodī-* ‘smell’); Arm. *gah*, *°owc* ‘throne, place’, *xrat*, *°owc* ‘judgement, advice’ and *mog*, *°owc* ‘wise man, Magian’

correspond respectively to OIr. **gāðu-* (OP *gāðu-*), **xratu-* (Av., OP *xratu-*) and **magu-* (OP *magu-*)¹.

Although exceptions to these correspondences are not uncommon (e.g. Arm. *ganj*, *owc* ‘treasure’ vs. OIr. **ganza-*)², all scholars agree that the coincidences are so many that they cannot be attributed to chance, and therefore require a proper explanation. The problem immediately stands out as a quite knotty one, since it involves several unknown factors such as the dating of the oldest Iranian borrowings into Armenian and the chronology of the loss of ancient final syllables both in Armenian and in W(estern)M(iddle)Ir(anian) dialects. Depending on which of these phenomena is chosen as a reference point, quite different explanations of the above-mentioned correspondences can be formulated. So, for example, Marquart (1911) – mostly ignoring the overtly MIr. phonology of the loanwords – assumed that the preservation of final vowels could only be attributed to an OIr. language, and consequently dated the borrowings to the Achaemenid period (6th-4th century BC). On the other hand, Meillet (1936), much more aware of the latest acquisitions concerning the historical phonology and dialectology of MIr. languages, knew that the source of most of the Iranian loanwords in Armenian must have been Parthian, which came into contact with the Armenians in the Arsacid period (1st-3rd century AD). Accordingly, he concluded that the correspondences between Iranian and Armenian stem classes should be regarded as evidence of the preservation of final syllables still at that time both in Parthian and in Armenian. Only afterwards, due to independent developments, Armenian (including Iranian loanwords) and Parthian would have lost their final syllables reaching the historically attested stage. Since Meillet’s view was the most widely accepted in the literature³, it is worth quoting it in full:

¹ Cf. MEILLET (1900: 258ff.), MARQUART (1911: 292), MEILLET (1911-1912: 17). Recently, a rich collection of examples has been offered by OLSEN (2005).

² Cf. again OLSEN (2005) for a complete list and see below for a discussion.

³ Cf. JENSEN (1959: 19), HUYSE (2003: 53-54) and BELARDI (2006: 34), essentially repeating Meillet’s formulation, and PERIKHANIAN (1966: 21-22 fn. 7), OLSEN (1999: 858-861; 2005) and KORN (2013), who share Meillet’s overall reasoning but propose to identify a different source dialect preserving the OIr. final syllables (respectively ‘Middle Median’, an Eastern Iranian dialect and Late Old Persian). On Olsen and Korn’s view see further below.

La chute des finales, qui a été plus complète que celle des autres voyelles, est elle-même postérieure aux anciens emprunts à l'iranien; en effet les thèmes nominaux iraniens en *-a-*, *-i-*, *-u-* donnent, dans ces vieux emprunts, des thèmes arméniens en *-a-* (ou en *-o-*), *-i-*, *-u-* [...]: au moment où ont été faits les emprunts, les mots pehlevi n'avaient donc pas encore perdu leur finale, et c'est en arménien que les finales sont tombées en même temps qu'elles tombaient aussi sur sol iranien; l'absence d'une finale *u* dans *xrat* en pehlevi et en arménien résulte de deux développements parallèles et indépendants. (Meillet, 1936: 23)

2. *Remarks on relative chronology*

Despite being very influential, Meillet's theory cannot be correct, as was brilliantly shown by Bolognesi (1954: 123-131). In this article, which has been given too little attention by the scholarly public⁴, the author demonstrates by means of relative chronology that the loss of final syllables in Armenian (the so-called *Auslautgesetz*) must have preceded the oldest Iranian borrowings. Consequently, the Iranian loanwords must have entered Armenian when they had already lost their final syllables. Let us only quote the most compelling among Bolognesi's arguments:

- The phenomena of metathesis in obstruent + *r* clusters (cf. Schmitt, 2007²: 72-73) and vowel prothesis before *r* (cf. Schmitt, 2007²: 77 and Lazzeroni, 1958) must predate the Iranian loanwords and follow the Armenian *Auslautgesetz*, according to the following chronology:
 - a. loss of final syllables except in monosyllables;
 - b. beginning of vowel prothesis before *r*;
 - c. metathesis of obstruent + *r* clusters in internal and initial position (e.g. PIE **b^hrātēr* > Pr.-Arm. **erbayr* > Arm. *etbayr*);
 - d. oldest layer of Iranian loanwords, showing vowel prothesis before *r* (e.g. Arm. *eramak* "herd", MP *ramag*) but no metathesis (Arm. *brinj* "rice", Parth. *brinz*);
 - e. end of vowel prothesis before *r*.

⁴ The only scholar who recognized the correctness of Bolognesi's arguments is R. SCHMITT (1983: 98-101; 2007²: 32-33).

Point (a) cannot be exchanged with point (b), because otherwise a monosyllable such as PIE **trins* “three” (Acc.Pl.) would have first become Pr.-Arm. **eris* and then Arm. ***ers* rather than the attested *eris* (cf. already Grammont, 1908: 235-236). In addition, point (c) cannot predate point (b), because an initial group *rC-* would have hardly developed without the support of a prothetic vowel (cf. Grammont, 1908: 234-238, followed by Schmitt, 2007²: 78).

- The change of **-w-* to *-g-* in initial and internal position but not in final position only affects native Armenian words (cf. Kim, 2021: 375-378). This means that when the oldest Iranian borrowings entered Armenian, both the loss of final vowels and the change of **-w-* to *-g-* had already happened. The chronology of these phenomena must have been the following:
 - a. loss of final vowels in Armenian except in monosyllables: PIE **g^wou-* > Arm. *kov* “cow” (cf. Martirosyan, 2010: 372); PIE **g^wou-ijo-* > Pr.-Arm. **kowi* “butter”; PIE **h₂reyi-* > Arm. *arew* “sun”, Gen.Sg. PIE **h₂reyi-os* > Pr.-Arm. **arewi* (cf. Martirosyan, 2010: 135-138)⁵;
 - b. change of **-w-* to *-g-* in initial and internal position: PIE **uorǵo-* > Pr.-Arm. **worc* > Arm. *gorc* “work”; Pr.-Arm. **kowi* > Arm. *kogi*; Pr.-Arm. **arewi* > Arm. **aregi*⁶; no change in *kov*, *arew* etc.;
 - c. oldest layer of Iranian borrowings, preserving *-w-* in all positions: Arm. *vkay* “witness” < Ir. **vikāvyā-* (Parth. *wigāh*, MP *gugāy*, see Tremblay, 2003: 132-133); *hrawēr* “invitation” < Ir. **fravāida-*⁷; Arm. *dew* < Ir. **daiva* (see above).

⁵ The reconstruction of the exact PIE declensional pattern to which the antecedents of Arm. *kov* and *arew/areg* belonged is to some extent conjectural, as Martirosyan himself recognizes. However, what is relevant to our discussion is that **-w-* was preserved in Armenian when followed by one single syllable (or no syllable) in the proto-form, i.e. when it became word-final after the action of the *Auslautgesetz*.

⁶ This form is indirectly attested in fixed expressions and compounds such as *areg-akn*, lit. “eye of the sun” and *Areg k’alak* “city of the sun, Ἡλίου πόλις”, cf. again MARTIROSYAN (2010: 135-138).

⁷ Cf. HÜBSCHMANN (1897: 183) and BOLOGNESI (1960: 41), who quotes this form to exemplify some of the phonological traits distinguishing the oldest Iranian borrowings.

Any other order of these three steps would contradict the attested forms: a sequence (b) → (c) → (a) would yield ***kog* instead of *kov* alongside *kogi*; a sequence (c) → (b) → (a) (assuming that the Iranian borrowings entered Armenian preserving their OIr. stem vowel) would yield ***kay*, ***deg* and ***bragēr* etc.⁸.

These arguments seem to me quite compelling, and therefore I do not see any way to avoid Bolognesi's conclusion:

Da quanto abbiamo detto dovrebbe essere chiaro che in armeno la sillaba finale è caduta prima degli imprestiti iranici; e ciò ci porta a stabilire che le parole iraniche sono passate in armeno quando già avevano a loro volta perduto la sillaba finale. (Bolognesi, 1954: 131)

As a consequence, all the attempts to solve the problem under examination postulating the preservation of the OIr. final syllables in the source language (see above fn. 3) cannot be accepted. However, one must admit that the alternative solution proposed by Bolognesi is far less convincing than his *pars destruens*. His idea, shared by Schmitt (1983: 98-101 and 2007²: 32-33), is that the choice of the Armenian stem class for Iranian borrowings was oriented by the comparison with derivatives and compounds in which the stem vowel was preserved. So, the stem vowel of Arm. *goyn* 'colour' (< Ir. **gauna-*) should have been inferred from the comparison with Arm. *gownak* (< Ir. **gaunaka-*), and the same should have happened with the couples Arm. *xrat* ~ *xratowk*, *mog* ~ **mogowpet* (later Arm. *mogpet*) etc. Ingenious as it may be, this position has some «severe drawbacks», as was rightly pointed out by Olsen (2005: 474-475). First, it is hardly conceivable that a so subtle analysis could spontaneously take place in a living language. I do not see why uncultivated speakers should have felt the need to pursue such a comparative

⁸ KIM (2021: 377 fn. 29) states that the change of internal and initial *-w-* to *-g-* must have happened after the apocope of final syllables and «thus after the bulk of Iranian borrowings entered Armenian» – so, an order (c) → (a) → (b) – referring to OLSEN (2005) and KORN (2013) «for evidence from nominal stem classes for survival of final vowels at this stage». This statement clearly disregards BOLOGNESI's (1954) conclusions and contradicts the data, because Ir. loanwords were never affected by the change *-w-* > *-g-* whereas, in this perspective, they should have been treated just like native Arm. words.

survey to select the declension for Iranian loanwords rather than inserting them systematically into the most productive nominal class. In addition, Olsen observed that this explanation would not work for the great number of suffixes borrowed from Iranian (Arm. *-ak* < Ir. **-aka-*; Arm. *-ik* < Ir. **-ika-*; Arm. *-owk* < Ir. **-uka-*; Arm. *-kert* < Ir. **-kṛta-*; Arm. *-stan* < Ir. **-stāna-*) which are almost never further derived or used as first members of compounds in Iranian, and yet are regularly assigned to the *-a-* declension in Armenian, as expected from their OIr. etymology.

In the end, it seems that all the explanations hitherto proposed for the puzzling coincidences between Armenian and OIr. stem classes are either unconvincing or simply impossible, so that it is worth trying to formulate a new solution to this problem.

3. *A different approach to Iranian-Armenian morphological interference*

As a premise to my proposal, I would like to draw attention on the theoretical framework usually adopted in discussing Iranian borrowings in Armenian. As should have appeared from the previous discussion, all scholars dealing with our matter assumed that the Iranian nouns entered Armenian in one single form of their paradigm (WMIr. Dir.Sg. according to most scholars, optionally OIr. or Eastern MIr. Acc.Sg. and neuter Nom./Acc.Sg. according to Olsen, 2005 and Korn, 2013). Admittedly, the selection of a single *Leitkasus* – as it was defined by Frei (1958: 2-3) – is a very common pattern by which morphological adaptation of loanwords takes place⁹. Obviously, in this perspective, the selection of the target stem class could only be based on that single *Leitform* so that denying that OIr. final vowels were still preserved at the time of the borrowing equals to denying the possibility that any information concerning the OIr. declension was available to the Armenian speakers exposed to the Iranian words. However, this is not the only possible scenario. Especially in the context of a very close linguistic contact, implying a high degree of bilingualism – as was surely the case

⁹ Cf. SCHMITT (1973: 12ff.) and GUSMANI (1986²: 45-50).

in Arsacid Armenia¹⁰ –, it is not uncommon that speakers are exposed to borrowed words in all the forms of their paradigm and not in one single *Leitkasus*¹¹. And it is precisely in a somewhat collateral form of the nominal paradigm that lies the possible solution to our problem.

4. *Western Middle-Iranian oblique plural case*

As is well known, the earliest documents of Middle Persian and Parthian still show a two-case morphological system, opposing a direct case to an oblique case¹². In the singular only personal pronouns and nouns of relationship (“father”, “mother” etc.) preserve a trace of the oblique case¹³, whereas in the plural it is always marked. The most common ending for the oblique plural is *-ān*, going back to the OIr. Gen.Pl. **-ānām*, but there are a number of instances in which the residual forms *-īn* and *-ūn* are attested, going back respectively to the OIr. Gen.Pl. of *-i-* and *-u-* stems (**-īnām* and **-ūnām*). These residual endings in some cases are remnants of the OIr. declension of the noun (e.g. MP *zanīn*, Parth. *žanīn* < OIr. **janīnām*, Av. *jaini-* “woman”), but more often are analogically extended to nouns or pronouns which in origin belonged to different inflectional classes¹⁴. Since we

¹⁰ Cf. MEYER (2017: 255-339), MANCINI (2008, citing the most relevant works by G. Bolognesi on the subject and offering personal observations on Armenian-Iranian contact) and BAILEY and SCHMITT (1986).

¹¹ Cf. SCHMITT (1973: 13) and especially GUSMANI (1986²: 57): «In fine, là dove due comunità linguistiche sono in stretto e durevole contatto, che si traduce in un frequente scambio di prestiti, si stabiliscono delle formule di corrispondenza fisse (analoghe a quelle che regolano i rapporti sul piano fonologico) che vengono automaticamente applicate ogni volta che si tratta di assimilare dal punto di vista morfologico un nuovo prestito».

¹² On the WMIr. two-case system cf. MANCINI (2019: 542-545 fnn. 28, 29, 31); DURKIN-MEISTERERST (2014: 197-201); CANTERA (2009); SKJÆRVØ (2009: 205); MANCINI (1992); SUNDERMANN (1989a: 130); SUNDERMANN (1989b: 154-155); SKJÆRVØ (1983); RASTORGUEVA and MOLČANOVA (1981: 188-189); SIMS-WILLIAMS (1981: 165-171); NYBERG (1974: 278); and lastly BARTHOLOMAE (1923: 9ff.), largely outdated but still useful.

¹³ E.g. Parth. *az* “I” vs. *man* “me”; WMIr. *pid* “father” Dir.Sg. vs. *pidar* Obl.Sg., see the literature in the previous footnote.

¹⁴ See the literature in fn. 12. The plural forms with an ending *-īn* were already mentioned by MEILLET (1911-1912: 17) in relation to the problem under examination, although, as we saw above, he did not take them into account in his later works on the subject.

are dealing with a small group of examples, it is worth quoting them in full (1-2)¹⁵:

(1) Obl.Pl. ending *-īn*:

- *ādurīn* “fires” (IPa <’tryn>): opposed to IMP *ādurān* (<’twr’ny>), originally a *-r-* stem **ātar/ātṛ-*.
- *Arminīn* “Armenians” (IPa <’rmny’n>): opposed to IMP *Arminān* (<’lmn’n>). A less likely reading is *Arminyān*, which would have probably been written <’rmny’n> (cf. <’ry’n> *Aryān* “Aryans”).
- *āwādīn* “generations” (Ps <’wb’tyn>), uncertain etymology¹⁶.
- *brādarīn* “brothers” (Ps <AHYtlyn>), originally a *-r-* stem **brātar-*.
- *dušmenīn* “enemies” (IPa <dwšmnyn>, MPa <dwšm(y)nyn>): originally a *-u-* stem **dušmanyu-* (Av. *dušmainiiu-*), see also below *dušmenūn*.
- *frawardīn*, name of a day and a month of the Zoroastrian calendar (IMP, IPa <prwrtyn>, IMP <plwltyn>, BP <plwltyn’>): the preservation of the OIr. Gen.Pl. **frawṛtīnām* (Av. *frawuuašīnqm*) is probably due to the fact that the word was perceived as a self-standing religious term rather than an inflected form of *fraward*.
- *frazandīn* “children” (IMP <plcndyn>, MPa <frzyndyn>): continuing the OIr. *-i-* declension (cf. Av. *frazainti-*), see also below *frazandūn*.

¹⁵ An (almost) complete collection of the attested forms can be found in DURKIN-MEISTERERNST (2014: 199-201). In the subsequent list the following abbreviations are used: I(nscriptional)M(iddle)P(ersian); (Pahlavi)Ps(alter); B(ook)P(ahlavi); M(anichean)M(iddle)P(ersian); I(nscriptional)Pa(rthian); M(anichean)Pa(rthian). Unless otherwise indicated, the words are quoted according to DURKIN-MEISTERERNST (2004) for Manichean texts, GIGNOUX (1972) for Sasanian inscriptions, MACKENZIE (1971) for Book Pahlavi, and SKJÆRVØ (1983: 163-176) for the Pahlavi Psalter.

¹⁶ Cf. SHAKI (1988: 95-96) who discusses previous literature and proposes a comparison with an OP form **uvādā-* “genealogy” traditionally restored in DB IV, 90-91. Despite being quite charming, this solution is hardly satisfying, because the sequence <[...]v-a-d-a> in DB IV, 90 could be read and restored otherwise (see FATTORI, forthc.), and even if the restoration were correct, the expected MP outcome would have been something like **xwāy* (Ir. **hv-* > OP *uv-* but MP *xw-* as in OP *uvaipašiya-*, MP *xwēbaš* and OP *-d-* > MP *-y-* as in OP *rādi*, MP *rāy*).

- *harwīn* “all” (<hrwyn> common in MPa, very rare in MMP; BP <hlwyn'>): originally a *-a-* stem (Av. *hauruuu-*).
- *harwispīn* “all” (BP <hlwspyn'>): originally a *-a-* stem (Av. Gen.Pl. *vīspanqm*).
- *hawīn* “they” (MPa <hwyn>, IPa <LHwyn>): analogical formation from the singular *hō* (< OIr. Nom.Sg. **hau*, OP *hau*, OAv. *huuō*, YAv. *hāu*)¹⁷. The original Gen.Pl. **avaišām* (OP *avaišām*, Av. *auuaēšqm*) is continued by MP *awēšān*.
- *imīn* “these” (MPa, MMP <'ymyn>): analogical formation from the singular *im* (< OIr. **ima-* found in OP Acc.Sg. *imam*, YAv. *iməm*). The original Gen.Pl. was **aišām* (Av. *aēšqm*).
- *kadagxwadāwīn* “houselords” (IPa <ktkhwtwyn>): most likely from a *-a-* stem **hvatāvya*¹⁸.
- *pusarīn* “sons” (IMP <BRrryn>, MMP <pwsryn>): originally a *-a-* stem (OP *puṣa-*, Av. *puṣra-*). The *-ar-* suffix is certainly extended from other nouns of relationship such as *pid*, *°arūn* (see below), *brād*, *°arīn* etc.
- *rad(a)nīn* “jewels” (MPa <rdnyn>): loanword from Skt. *ratna-* “jewel”.
- *wxārīn* “sisters” (MPa <wx'ryn>) and *xwārīn* “id.” (MMP <xw'ryn>): originally a *-r-* stem **hvahar-*. The Obl. stem *wxār-/xwār-* to which the ending *-īn* is attached could trace back to the OIr. stem **hvahr-* with pre-desinential zero grade (cf. OAv. Gen.Pl. *dugədrqm*, YAv. Gen.Pl. *duγδrqm* “of daughters”)¹⁹, whereas in the other nouns of relationship the full-grade stem was analogically extended²⁰.

¹⁷ Cf. RASTORGUEVA and MOLČANOVA (1981: 205).

¹⁸ I find a derivation from **hvatāvya-* (cf. BENVENISTE, 1966: 29-30 with further literature) more convincing than **hvatāvan-*, often mentioned as an alternative (see e.g. BARTHOLOMAE, 1920 *passim*), since it would explain better the orthography with final <wy> in IPa and the palatal umlaut in the last syllable of Bactr. *χoαδηo*.

¹⁹ Cf. BARTHOLOMAE (1923: 4-5).

²⁰ MP and Parth. Obl.Sg. *pidar* < **pitarahya* (analogically built on Acc.Sg. **pitaram*) rather than from OIr. Gen.Sg. **piṣrah* which would have yielded MP **pis* (cf. OP Gen.Sg. *piṣa*) and Parth. **pihr* (cf. Av. Gen.Sg. *piṣrō*), cf. CANTERA (2009: 28-29). On the singular form BP *xwāhar/xwahār* cf. CANTERA (2009: 25 fn. 24).

- *xšahrdārīn* “rulers” (IPa <xštrdryn>): also attested as *xšahrdārān* (IPa <xštrdrn>, cf. IMP *šahryārān* <štrd’r’n>). Originally a *-a-* stem **xšaδradāra-*.
- *žam(a)nīn* “times, hours” (MPa <jmnyn>): it is not clear whether this form should be regarded as a different OIr. formation than <jm’n> *žamān* “time” (as BP *zaman* “hour” opposed to *zamān* “time”, cf. Panaino, 2017: 154) or as its Obl.Pl. case built attaching (non-etymological) *-īn* to an ancient oblique stem. This last option would support the idea that this noun was originally an ablauting *-n-* stem **jaman-* (Acc.Sg. **jamānam*, Gen.Sg. **jamanah* etc.) as suggested by Panaino (2017: 167-172).
- *žanīn* “women” (MPa <jnyn>) and *zanīn* “women” (MMP <znyn>): these forms clearly continue the OIr. Gen.Pl. **janīnām* (Av. *jaininām*).

(2) Obl.Pl. ending *-ūn*:

- *ardāyūn* “righteous ones” (Ps <’lt’dwny>): originally a *-n-* stem, cf. OP *rtāvan-*.
- *dušmenūn* “enemies” (MMP, MPa <dwšmnwn>): this form probably continues the OIr. Gen.Pl. **dušmanyūnām* (Av. *duš-mainiunām*).
- *pidarūn* “fathers” (Ps <ptlwny>): originally a *-r-* stem **pitar-*.
- *frazandūn* “children” (Ps <plcndwn>): originally a *-i-* stem, see above.
- *mowūn* “Magi” (IMP <mgwny>): the form continues the OIr. Gen.Pl. **magūnām* (OP *magu-*).

5. *A novel solution to an old problem*

It seems reasonable to assume that the scenario attested in the Sasanian period was the last stage of a long-term process of substitution of the older endings *-īn* and *-ūn* with a general Obl.Pl. ending *-ān*, which was eventually going to displace them completely. This implies that a previous stage existed when the distribution of the endings *-īn* and *-ūn* was more consistent with the OIr. stem classes and less influenced by

analogy. If one considers that the earliest examples quoted above date back to the beginning of the Sasanian period (mid-3rd century AD), i.e. approximately two centuries after the establishment of the Arsacid rule over Armenia (mid-1st century AD)²¹, it does not seem unlikely that at the time of the oldest Iranian borrowings into Armenian the WMIr. Obl.Pl. endings still reflected the OIr. stem vowels to a great extent.

At a closer look, the nominal paradigms in pre-Sasanian WMIr. as they have just been described clearly appear to be very similar to the Armenian ones from a structural point of view. Both have a singular ‘direct’ case form (= Arm. Nom./Acc.Sg.) unmarked with respect to declension and a plural ‘oblique’ case form (= Arm. Gen./Dat./Abl.Pl.) which bears an unambiguous mark of the noun’s inflectional class²², as shown in the following scheme (Table 1).

	Pre-Sasanian WMIr. (Parthian)		Armenian	
	Dir.Sg.	Obl.Pl.	Nom./Acc.Sg.	Gen./Dat./Abl.Pl.
<i>-a-</i> stems	* <i>dāt</i> “law”	* <i>dātān</i>	<i>am</i> “year”	<i>amac’</i>
<i>-i-</i> stems	* <i>zan</i> “woman”	* <i>zanīn</i>	<i>ban</i> “word”	<i>banic’</i>
<i>-u-</i> stems	* <i>mog</i> “Magian”	* <i>mogūn</i>	<i>cov</i> “sea”	<i>covowc’</i>

Table 1. *Comparison between WMIr. and Arm. vowel stem paradigms.*

Thus, there is no difficulty in assuming that this comparability between nominal paradigms could grant the correspondence between Iranian and Armenian stem classes (WMIr. **mog*, **ūn* → Arm. *mog*, **owc’* etc.), without having to resort to the preservation of the OIr. final syllables or to other complex explanations.

²¹ Cf. SCHMITT (1983: 73-74).

²² Although Armenian had a more complex nominal morphology than WMIr., the Gen./Dat./Abl.Pl. was surely the most representative form with regard to the word’s declension, since the other case endings were either unmarked (Nom./Acc.Sg. -ø, Nom.Pl. -*k’*; Acc./Loc.Pl. -*s* for all vowel stems), ambiguous (Gen./Dat./Loc. -*i* both for *-a-* and *-i-* stems; Abl.Sg. -*ē* for *-a-*, *-i-* and *-ow* stems) or less frequently used (Instr.Sg. and Pl.). This special status of the Gen./Dat./Abl.Pl. is well exemplified by the modern lexicographical practice of providing this case form (rather than the Gen.Sg. as in Greek and Latin) beside the Nom./Acc.Sg. to synthetically indicate a noun’s declension.

In addition, the frequent instances of analogical extension of *-īn* and *-ūn* to nouns originally belonging to other inflectional classes in WMIr. represent a convincing starting point to explain the analogous mismatches attested in Armenian²³. Even if, as was claimed above, it is likely that in pre-Sasanian WMIr. the distribution of the Obl.Pl. endings *-īn* and *-ūn* was more consistent with the OIr. stem vowels, it cannot be excluded that analogy was already operating. Of special significance in this respect are the examples of extension of the ending *-ūn* listed above, since, as was pointed out by Olsen (2005: 477-478), a shift towards the *-u-* declension would be difficult to justify as an inner Armenian development²⁴.

If one accepts the proposed explanation, the preservation of the OIr. stem classes in Armenian does not only cease to be an unsolved problem, but also becomes a significant piece of linguistic evidence. On the one hand, it sheds light on a hitherto underestimated aspect of the Armenian-Iranian linguistic contact, i.e. the role played by the strong similarity of the WMIr. and Armenian morphological systems in favoring lexical borrowing²⁵. On the other hand, it represents a contribution of some weight to the reconstruction of the historical development of Western Iranian morphology from the Old to the Middle Iranian period, which is notoriously very difficult to research on due to lack of direct sources.

²³ This was already realized by MEILLET (1911-1912: 17): «mais déjà ces anciens pluriels n'étaient plus compris [...]. On conçoit donc que les emprunts arméniens n'aient pas tous exactement la forme des thèmes attendus».

²⁴ A possible objection to this is that the only extant examples of non-etymological extension of *-ūn* are found in Middle Persian, whereas the main source of Iranian loanwords in Armenian is Parthian. However, the isolated attestations of *ardāyūn*, *pidarūn* and *frazandūn* in the Pahlavi Psalter are hardly enough to assume that the analogical extension of *-ūn* was a typically Middle Persian feature. The lack of evidence for *-ūn* in Parthian could simply be due to our scanty documentation.

²⁵ Incidentally, one may observe that Armenian and WMIr. are also quite similar with respect to their verbal inflection – especially in the present indicative – and this is probably one of the reasons behind the great inclination shown by Armenian to borrow not only nouns and adjectives but also a significant number of Iranian verbs.

References

- BAILEY, H.W. and SCHMITT, R. (1986), *Armenia and Iran IV. Iranian influences in Armenian Language*, in YARSHATER, E. (1986, ed.), *Encyclopaedia Iranica*. Vol. 2, 4-5, Routledge & Kegan Paul, London, pp. 445-465.
- BARTHOLOMAE, CH. (1920), *Zur Kenntnis der mittelliranischen Mundarten III*, Carl Winter, Heidelberg.
- BARTHOLOMAE, CH. (1923), *Zur Kenntnis der mittelliranischen Mundarten V*, Carl Winter, Heidelberg.
- BELARDI, W. (2006), *Elementi di armeno aureo*. Vol. 2: *Le origini indoeuropee del sistema fonologico dell'armeno aureo*, Il Calamo, Roma.
- BENVENISTE, É. (1966), *Titres et noms propres en iranien ancien*, Klincksieck, Paris.
- BOLOGNESI, G. (1954), *Ricerche sulla fonetica armena*, in «Ricerche Linguistiche», 3, pp. 123-154.
- BOLOGNESI, G. (1960), *Le fonti dialettali degli prestiti iranici in armeno*, Vita e Pensiero, Milano.
- CANTERA, A. (2009), *On the history of the Middle Persian nominal inflection*, in SUNDERMANN, W., HINTZE, A. and DE BLOIS, F. (2009, eds.), *Exegisti Monumenta. Festschrift in Honour of Nicholas Sims-Williams*, Harrassowitz, Wiesbaden, pp. 17-30.
- DURKIN-MEISTERERNST, D. (2004), *Dictionary of Manichaean Middle Persian and Parthian*, Brepols, Turnhout.
- DURKIN-MEISTERERNST, D. (2014), *Grammatik des Westmittelliranischen (Parthisch und Mittelpersisch)*, Österreichische Akademie der Wissenschaften, Wien.
- FATTORI, M. (forthc.), *New Epigraphic and Exegetical Remarks on Paragraph 70 of the Bisotun Inscription, Old Persian Version (DB/OP IV ll. 88-92)*.
- FREI, P. (1958), *Die Flexion griechischer Namen der I. Deklination im Latein*, Keller, Winterthur.
- GIGNOUX, P. (1972), *Glossaire des inscriptions pehlevies et parthes*, Lund Humphries, London.

- GRAMMONT, M. (1908), *La métatèse en arménien*, in *Mélanges de linguistique offerts à M. Ferdinand de Saussure*, Champion, Paris, pp. 231-243.
- GUSMANI, R. (1986²), *Saggi sull'interferenza linguistica. Seconda edizione accresciuta*, Le Lettere, Firenze.
- HÜBSCHMANN, H. (1897), *Armenische Grammatik*, Breitkopf & Härtel, Leipzig.
- HUYSE, P. (2003), *Le y final dans les inscriptions moyen-perses et la 'loi rythmique' proto-moyen-pers*, Association pour l'avancement des études iraniennes, Paris.
- JENSEN, H. (1959), *Altarmenische Grammatik*, Carl Winter, Heidelberg.
- KIM, R. (2021), *The Phonetics and Phonology of Old Armenian* <V>, <W>, and Prevocalic <OW>, in «Transactions of the Philological Society», 119, 3, pp. 371-392.
- KORN, A. (2013), *Final troubles: Armenian stem classes and the word-end in Late Old Persian*, in TOKHTASJEV, S. and LURJE, P.B. (2013, eds.), *Commentationes Iranicae. Sbornik statej k 90-letiju Vladimira Aronoviča Livšica*, Nestor-Istorija, St. Petersburg, pp. 74-91.
- LAZZERONI, R. (1958), *Ipotesi sulla vocale protetica davanti a -r- in greco e in armeno*, in «Annali della Scuola Normale Superiore di Pisa. Lettere, Storia e Filosofia», 27, 2, pp. 127-136.
- MACKENZIE, D.N. (1971), *A Concise Pahlavi Dictionary*, Oxford University Press, London / New York / Toronto.
- MANCINI, M. (1992), *Una nuova testimonianza sul caso obliquo tra persiano antico e mediopersiano*, Istituto di scienze storico-filologiche, Viterbo.
- MANCINI, M. (2008), *Contatto e interferenza di lingue nei lavori orientalisti di G. Bolognesi*, in FINAZZI, R.B. e TORNAGHI, P. (2008, a cura di), *Dall'Oriente all'Occidente. Itinerari linguistici di Giancarlo Bolognesi, Giornata di studio (4.5.2007)*, Università Cattolica del Sacro Cuore, Milano, pp. 23-52.
- MANCINI, M. (2019), *Middle-Persian Morphology and Old Persian Masks. Some Reflections on 'Proto-Middle Persian'*, in BADALKHAN, S., BASELLO, G.P. and DE CHIARA, M. (2019, eds.), *Iranian Studies in Honour of Adriano Valerio Rossi*. Vol. 2, UniorPress, Napoli, pp. 523-565.

- MARQUART, J. (1911), *Armenische Streifen*, in AKINIAN, P.N. (1911, Hrsg.), Hushardzan. *Festschrift aus Anlass des 100 jährigen Bestandes der Mechitharisten-kongregation in Wien (1811-1911) und des fünfundzwanzigsten Jahrganges der philologischen Monatsschrift ‚Handes Amsorya‘ (1887-1911)*, Mechitharisten-Buchdruckerei, Wien, pp. 291-302.
- MARTIROSYAN, H.K. (2010), *Etymological Dictionary of the Armenian Inherited Lexicon*, Brill, Leiden / Boston.
- MEILLET, A. (1900), *La déclinaison et l'accent d'intensité en perse*, in «Journal asiatique», 15, pp. 254-277.
- MEILLET, A. (1911-1912), *Sur les mots iraniens empruntés par l'arménien*, in «Mémoires de la Société de Linguistique de Paris», 17, pp. 242-250.
- MEILLET, A. (1936), *Esquisse d'une grammaire comparée de l'arménien classique*, Imprimerie des PP. Mekhitharistes, Vienne.
- MEYER, R. (2017), *Iranian-Armenian language contact in and before the 5th century CE*, PhD thesis, University of Oxford, Oxford.
- NYBERG, H.S. (1974), *A Manual of Pahlavi*. Vol. 2, Harrassowitz, Wiesbaden.
- OLSEN, B.A. (1999), *The Noun in Biblical Armenian*, De Gruyter, Berlin / Boston.
- OLSEN, B.A. (2005), *On Iranian Dialectal Diversity in Armenian*, in MEISER, G. and HACKSTEIN, O. (2005, Hrsg.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft (17.-23. September 2000, Halle an der Saale)*, Reichert, Wiesbaden, pp. 473-481.
- PANAINO, A. (2017), *The origins of Middle Persian Zamān and related words: A controversial etymological history*, in «Iran and the Caucasus», 21, 2, pp. 150-195.
- PERIKHANIAN, A. (1966), *Une inscription araméenne du roi Artasēs trouvée à Zanguézour (Siwnik)*, «Revue des Études Arméniennes – N.S.», 3, pp. 17-29.
- RASTORGUEVA, V.S. and MOLČANOVA, E.K. (1981), *Parfianskij jazyk*, in RASTORGUEVA, V.S., ABAEV, V.A. and BOGOLJUBOV, M.N. (1981, eds.), *Osnovy iranskogo jazykoznanija*. Vol. 2: *Sredneiranskije jazyki*, Nauka, Moskva, pp. 147-232.

- SCHMITT, R. (1973), *Probleme der Eingliederung fremden Sprachgutes in das grammatische System einer Sprache*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.
- SCHMITT, R. (1983), *Iranisches Lehngut im Armenischen*, in «Revue des Études Arméniennes – N.S.», 17, pp. 73-112.
- SCHMITT, R. (2007²), *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen. 2. durchgesehene Auflage*, Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.
- SHAKI, M. (1988), *Pahlavica*, in SUNDERMANN, W., DUCHESNE-GUILLEMIN, J. and VAHMAN, F. (1988, eds.), *A Green Leaf: Papers in Honour of Professor Jes P. Asmussen*, Brill, Leiden / Téhéran / Liège, pp. 93-99.
- SIMS-WILLIAMS, N. (1981), *Notes on Manichaean Middle Persian Morphology*, in «Studia Iranica», 10, pp. 165-176.
- SKJÆRVØ, P.O. (1983), *Case in Inscriptional Middle Persian, Inscriptional Parthian and the Pahlavi Psalter*, in «Studia Iranica», 12, 2, pp. 47-62, 151-181.
- SUNDERMANN, W. (1989a), *Mittelpersisch*, in SCHMITT, R. (1989, ed.), *Compendium Linguarum Iranicarum*, Reichert, Wiesbaden, pp. 138-164.
- SUNDERMANN, W. (1989b), *Parthisch*, in SCHMITT, R. (1989, ed.), *Compendium Linguarum Iranicarum*, Reichert, Wiesbaden, pp. 114-137.
- TREMBLAY, X. (2003), *La résurrection du bactrien: à propos des "Bactrian Documents"*, in «Indo-Iranian Journal», 46, 2, pp. 119-133.

MARCO FATTORI

Dipartimento di Lettere e Culture Moderne

Università di Roma 'La Sapienza'

Piazzale Aldo Moro 5

00185 Roma (Italy)

marco_fattori@live.it



Sulla polifunzionalità di *vienas* “uno” in lituano antico

ADRIANO CERRI

ABSTRACT

This study aims to describe from a synchronic and typological perspective the variety of functions displayed by the word *vienas* “one” in Old Lithuanian. To do so, we examine the occurrences of this lexeme in eight of the oldest linguistic monuments of this language, all of which date from the mid- to late 16th century. In addition to the numeral function, *vienas* fulfils several other functions: nominal, pronoun, adjective and pseudo-article. The analysis of the last function requires both philological-textual and linguistic-typological considerations. The former allows us to note that the use of the lexeme ‘one’ is documented in existential / presentative constructions and it is favoured by (i) the presence of a similar lexical element in the source-text from which the Lithuanian texts were translated, and by (ii) stylistic features that bring the text closer to the folk-tale genre (where ‘one’ as presentative marker is more frequent than elsewhere). Typological considerations lead to the conclusion that ‘one’ in Old Lithuanian presents a limited degree of grammaticalisation, since it can serve as a presentative marker and, in some cases, as a specific marker (always, however, alternating with the much more frequent absence of overt marking).

KEYWORDS: one, grammaticalisation, indefiniteness, existential constructions.

1. *Introduzione*

È noto che in molte lingue il lessema indicante ‘uno’ svolge, oltre a quella primaria di numerale cardinale, anche altre funzioni secondarie, andando così a collocarsi in diverse categorie lessicali. La polifunzionalità del primo numerale cardinale è un fenomeno frequente, osservabile in numerose lingue del mondo e presente anche in lituano¹.

¹ Si rileva che l’interesse dei linguisti nei confronti di ‘uno’ si è concentrato principalmente sui processi di grammaticalizzazione osservabili in molte lingue antiche e moderne (cfr. GIVÓN, 1981; HEINE, 1997: 71ss.; MEISTERFELD, 2000; DRYER, 2013). La funzione

Il principale strumento lessicografico di questa lingua, il *Dizionario della lingua lituana* (*Lietuvių kalbos žodynas*), individua tre diverse entrate per il lessema *vienas*, -à: (i) numerale cardinale; (ii) pronome; (iii) aggettivo (cfr. LKŽe, s.v.).

Questa varietà di funzioni non è, ovviamente, una peculiarità del lituano contemporaneo; al contrario, già nei testi antichi si riconoscono diverse funzioni grammaticali di ‘uno’ (cfr. ALEW II: 1233-1236). La ricerca delle ricorrenze di *vienas* (<wienas>, <wenas>, <weenas>, <wiens> e altre varianti grafiche) nei monumenti linguistici lituani antichi rivela una casistica piuttosto complessa all’interno della quale, oltre alla funzione propriamente numerale (ovvero la *numerazione*), il lessema sotto indagine risponde ad altre funzioni non-numerali. In mancanza di uno studio specifico su questo punto, il presente lavoro si propone di analizzare un ampio numero di ricorrenze del lessema ‘uno’ tratte da un corpus di otto testi lituani del XVI secolo al fine di descriverne la polifunzionalità².

Sebbene le due nozioni siano strettamente correlate, si fa qui riferimento alle funzioni grammaticali anziché ai significati; di conseguenza si parlerà di polifunzionalità anziché di polisemia. A rigore, una situazione in cui un elemento linguistico mostra più di un significato dovrebbe rientrare nel dominio della polisemia, ma è tuttora oggetto di dibattito se la varietà di significati vada riconosciuta come una proprietà intrinseca del lessema (che in questo caso sarebbe polisemico), o se, al contrario, sia più corretto ricondurre i vari significati a un unico

non-numerale più studiata è quella di articolo indefinito (nelle lingue che lo hanno sviluppato o che lo stanno sviluppando). Particolarmente interessanti da questo punto di vista sono le lingue slave, nelle quali il fenomeno è relativamente recente o addirittura in fase incipiente, si veda per il bulgaro GEIST (2013) e per il polacco MOLINARI (2022). Nonostante la vasta bibliografia sul tema, le altre funzioni secondarie (non-numerali) di ‘uno’ restano assai meno studiate. Due valide eccezioni sono LUJÁN MARTINEZ (1995) e KUTEVA *et al.* (2019: 298-306) cui si farà riferimento più volte nel corso di questo lavoro.

² Gli esempi qui analizzati sono tratti, in quanto particolarmente rappresentativi, dal corpus di testi riportato in appendice. Le fonti prese in considerazione coprono l’arco temporale della prima stagione della scrittura lituana (1547-1599). Anche per un lasso di tempo così limitato i testi offrono un grandissimo numero di ricorrenze del lessema ‘uno’; basti pensare che una ricerca delle concordanze della radice <wien-> nella sola *Postilla* di Daukša (1599) produce 1134 risultati (<https://seniejirastai.lki.lt/db.php?source=2>; cfr. anche KUDZINOWSKI, 1977, ed.: 422-427).

significato di base (considerando così il lessema monosemico)³. Posta la questione in questi termini, sembra corretto considerare 'uno' un lessema monosemico polifunzionale, il cui nucleo semantico, o significato di base, è quello di 'singolarità'.

2. *Inquadramento teorico*

2.1. *Numerazione e quantificazione*

In generale, per i numerali polifunzionali – e a maggior ragione per 'uno' che, fra tutti, è quello che si presta alla massima varietà di funzioni – una prima distinzione oppone gli usi *numerali* a quelli *non-numerali*. I primi sono quelli in cui si manifesta la funzione della numerazione. Con essa si intende un modulo specifico della quantificazione la cui funzione caratteristica è individuare quantità assolute⁴. La classe lessicale che assolve a questa (sotto-)funzione è appunto quella dei numerali, ovvero lessemi (semplici o complessi) che hanno per referente un *numero* e che possono quantificare un N (cfr. McGregor, 2014: 50). La numerazione si distingue tanto dalla quantificazione relativa quanto da quella indeterminata. I mezzi attraverso cui si esprimono queste ultime due sotto-funzioni sono chiamati perciò quantificatori (e come tali si oppongono ai numerali); sono esempi di quantificatori del primo tipo le locuzioni italiane *meno di*, *più di*, *il quadruplo di*, ecc.; mentre sono quantificatori indeterminati *poco*, *tanto*, *alcuni*, *un po' di*, ecc. (cfr. Cardinaletti e Giusti, 2018).

La funzione prototipica dei numerali è, naturalmente, quella della numerazione, tuttavia alcuni di essi assolvono anche ad altre funzioni, prima tra tutte quella della quantificazione indeterminata; è il caso, ad esempio, dell'impiego approssimativo di it. *due*, *quattro* o *mille* in espressioni come *scambiare due parole*, *fare quattro risate*, *ripetere mille volte*, ecc. (cfr.

³ Su questo punto ci si rifà alle proposte di WHERRITY (2016) a proposito di un approccio integrato tra monosemia e polisemia.

⁴ Vedi la distinzione operata da DE LA VILLA (2010: 175-176), relativamente ai numerali latini, tra *absolute numerals*, ovvero quelli indicanti quantità assolute (*unus*, *duo*, *decem*, ecc.), e *relative quantities* (*aliqui*, *pauci*, *multi*, *omnes*, ecc.).

Krifka, 2007; Bazzanella, 2011). In questi casi i lessemi numerali sono impiegati come quantificatori. Preme sottolineare che, differenza di tutti gli altri numerali, ‘uno’ non si presta a passare da numerale a quantificatore a causa della sua singolarità intrinseca: l’unità non può mai fungere come pluralità indistinta⁵. Quindi se per i numerali diversi da ‘uno’ l’opposizione tra funzione numerale e non-numerale corrisponde a quella tra numerazione e quantificazione indeterminata, per ‘uno’ le funzioni non-numerali vanno ricercate al di fuori dell’ambito della quantificazione.

2.2. Funzioni numerali e non-numerali

Secondo Wiese (2003: 9-42; 2007) si possono distinguere tre tipi di utilizzo dei numerali: (i) cardinale; (ii) ordinale; (iii) nominale. Nel primo caso il numerale indica una quantità o cardinalità (cioè l’estensione di un insieme di elementi numerabili), p. es. *cinque autobus*; nel secondo la posizione in una successione ordinata, p. es. *il quinto autobus*; nel terzo, infine, il numerale è semplicemente un’etichetta nominale, p. es. *l’autobus (numero) cinque*. Quest’ultima funzione è talvolta confusa con la seconda, ma differisce da essa in quanto non presuppone una successione ordinata: in questo caso *cinque* è solo un modo per identificare un autobus rispetto agli altri; dunque il valore del numerale consiste tutto nella sua funzione distintiva che lo oppone alle etichette nominali assegnate agli altri elementi del sistema (gli altri autobus). A livello formale, le funzioni (i) e (iii) sono svolte dai numerali cardinali, mentre la seconda dagli ordinali⁶.

⁵ Tuttavia sono noti casi in cui ‘uno’ modifica un altro numerale assegnando ad esso un valore approssimativo, p. es. tamil *oru* “uno” vale “circa” se impiegato come modificatore di un altro numerale: *oru ampatu* “circa cinquanta” (cfr. KUTEVA *et al.* 2019: 304-305).

⁶ Conviene però notare che non di rado negli usi concreti della lingua si registrano incoerenze quali la presenza di numerali cardinali in contesti ordinali, p. es. it. *il ventinove (di...)*, sp. *el veintinueve (de...)* per indicare il *ventinovesimo* giorno di un mese (come testimoniano, fra le altre lingue, il ted. *der neunundzwanzigste...* e il ru. *двадцать девятое...*). In alcune lingue e in determinati ambiti (solitamente quelli che presentano degli elementi rigidamente strutturati in successione, p. es. le ore del giorno, i capitoli di un libro *et sim.*) l’uso dei numerali cardinali e ordinali può alternarsi sinonimicamente, p. es. lit. *pusė penktos (valandos)*, letteralm. “metà della quinta (ora)” alterna con *pusė penkių (valandų)* “metà delle cinque (ore)” per esprimere lo stesso orario (“quattro e mezza”). Per inciso, il fatto che le lingue attestino sconfinamenti delle forme cardinali nel dominio degli ordinali, mentre non mi risulta che sia dato il contrario, testimonia la priorità e non-marcatezza dei primi rispetto ai secondi.

Nonostante la funzione nominale del tipo *l'autobus cinque* faccia ricorso ai numerali cardinali, è evidente che essa non rientra nella numerazione (né più in generale nella quantificazione), perciò è più appropriato collocarla tra le funzioni non-numerali⁷. Ne consegue che l'unica funzione propriamente numerale di 'uno' è quella cardinale (v. § 3.1; quella ordinale notoriamente viene assolta da una forma distinta, derivata o ancora più spesso suppletiva).

Quanto alle funzioni non-numerali, la più nota e studiata (v. nota 1) è quella di articolo indefinito. Poiché il lituano è una lingua senza articoli, è utile richiamare un'osservazione di Luján Martínez (1995: 218): «incluso en lenguas en las que no existe gramaticamente dicha categoría [l'articolo indefinito] el numeral 'uno' presenta usos totalmente afines a los que ofrece el artículo indeterminado en el caso de las lenguas que sí que cuentan con ella». A sostegno di questa affermazione, l'autore presenta due esempi dal greco antico e dal latino in cui *εἷς* e *ūnus* si comportano in maniera molto simile ad articoli. Si vedrà che anche il lituano antico offre alcune testimonianze di questo stesso tipo (v. § 3.5). Altre funzioni non-numerali di 'uno' descritte da Luján Martínez (1995: 219-222) sono quella pronominale (v. § 3.3; cfr. anche Kuteva *et al.* 2019: 301) e quella anaforica (*pro-form*) tipica dell'inglese (che invece è assente in lituano):

- (1) *I have bought a new refrigerator and I don't know what to do with the old one.*

Conviene precisare che la peculiarità di questa costruzione non consiste solo nel riferimento anaforico (il quale si ritrova anche in altre costruzioni, p. es. in quelle correlative: ingl. *I have two refrigerators. One is new, the other is old*; cfr. lat. *Erant omnino itinera duo [...]: unum per Sequanos, [...] alterum per provinciam nostram*, Caes. *De bello gallico* I.6), ma anche nell'individuazione di un elemento rispetto agli altri. L'elemento

⁷ Ci si può invece domandare se casi di questo tipo non vadano analizzati nei termini di una relativa restrittiva ridotta: *l'autobus (numero) cinque* = *l'autobus (che è / ha il numero) cinque*. Visto in questa prospettiva, il lessema numerale svolgerebbe una funzione determinativa (modificazione). L'osservazione mi è stata suggerita da uno dei revisori anonimi, che ringrazio per questa come per tutte le altre preziose indicazioni.

così individuato assume automaticamente il tratto [+ definito]. Questo spiega la ragione per cui il lituano (sia antico che moderno) non utilizza *vienas* in costrutti del tipo (1), bensì ricorre alla forma pronominalizzata dell'aggettivo, ovvero quella che tipicamente veicola la definitezza (2):

- (2) *Nusipirkau nauja šaldytuva ir*
 comprare-PST.1.SG nuovo-ACC.SG frigorifero-ACC.SG e
nežinau, ką daryti
 NEG-sapere-PRS.1.SG cosa-ACC.SG fare-INF
su senuoju.
 con vecchio-INS.SG.DEF
 “Ho comprato un nuovo frigorifero e non so che cosa fare con
 il/quello vecchio.”

Oltre a quelle richiamate fin qui, i dati del lituano permettono di individuare la funzione aggettivale (v. § 3.4)⁸.

3. ‘Uno’ in lituano antico

3.1. Funzione cardinale

Tra i molti possibili esempi della funzione cardinale si riporta un passo tratto dal primo libro a stampa lituano, il catechismo luterano *Le semplici parole* di M. Mažvydas (1547)⁹:

- (3) *Balsinės venai dwem alba tryms*
 vocale-NOM.PL.F 1-DAT.F 2-DAT oppure 3-DAT
įąbalsinems pirm detas
 consonante-DAT.PL.F prima porre-PTCP.PST.PASS.NOM.PL.F
 “Vocali poste prima di una, due o tre consonanti.” (MžK 1547: 15₂₂)¹⁰

⁸ Questa non è presa in considerazione da LUJÁN MARTINEZ (1995); vedi invece KUTEVA *et al.* (2019: 298-299, 302).

⁹ Su Martynas Mažvydas e il primo libro lituano cfr. DINI (1997: 295, 414-416), KAUNAS *et al.* (1998). Più in generale sui monumenti linguistici lituani del XVI-XVII secolo cfr. ZINKEVIČIUS (1996: 227-255).

¹⁰ Le fonti sono indicate in questo modo: sigla (eventuale numero di tomo in numeri romani), anno, numero di pagina (eventuale indicazione *r* = *recto* o *v* = *verso*); infine, in

Da una verifica di tutte le ricorrenze di 'uno' negli scritti di questo autore, così come elencate nel lessico degli scritti di Mažvydas compilato da Urbas (1998: 418-419), si riscontra che, a dispetto dei 17 usi numerali ivi elencati, solo in due casi si può dire che 'uno' svolga questa funzione. Oltre a quello appena citato, l'altro esempio è il seguente:

- (4) *ščimty ųžmaniu wena negaleczia*
 100-LOC persona-GEN.PL 1-ACC NEG-potere-COND.PRS.1.SG
atrafti Kurfai wena ųžadi dewa
 trovare-INF REL 1-ACC.M parola-ACC.SG.M Dio-GEN.SG.M
prįsakima makietu Ir pateriaus
 comandamento-GEN.SG.M sapere-SUBJ.PRS.3 e preghiera-GEN.SG.M
ben du ųžadžiu atmintu
 almeno 2-ACC parola-ACC.DU ricordare-SUBJ.PRS.3
 "Tra cento persone non ne potrei trovare una che sappia una parola
 di un comandamento di Dio e che ricordi almeno due parole di una
 preghiera." (MžK 1547: 10₁₈₋₂₀)

Si può notare che, delle due ricorrenze di *wena*, solamente la seconda, che quantifica *ųžadi* "parola", ha funzione numerale (così come *du* nel successivo *du ųžadžiu* "due parole"); nel primo caso, invece, *wena* è un pronome anaforico in una costruzione partitiva: "una [persona] tra cento persone". In tutte le altre ricorrenze rubricate da Urbas sotto la voce 'numurale' ci troviamo piuttosto di fronte a uno dei vari usi non-numerali esemplificati più sotto.

Altri esempi di funzione numerale di 'uno' presso altri autori di epoca lituana antica sono i seguenti:

- (5) *wenai karaliųtai nepigai ųsdera*
 1-DAT.F regno-DAT.SG.F NEG-facile-ADV confare-PRS.3
tris karelius
 3-NOM sovrano-NOM.PL?
 "A un regno mal si addicono tre sovrani." (WP 1573: 50v₂₁₋₂₂)

pedice, numero di rigo/i. Per i passi biblici viene riportato anche il riferimento secondo le abbreviazioni italiane correnti (p. es. Gen = Genesi; At = Atti degli apostoli, ecc.). Tutte le traduzioni italiane dei passi analizzati sono mie.

- (6) *Wiēnas yrā Diéwas / wiēnas*
 1-NOM.M essere-PRS.3 Dio-NOM.SG.M. 1-NOM.M
krikštas / wiená bažnicžia /
 battesimo-NOM.SG.M 1-NOM.F chiesa-NOM.SG.F
wiēnas kūnas Chřistaus
 NOM.M corpo-NOM.SG.M Cristo-GEN
 “Uno è Dio, uno il battesimo, una la chiesa, uno il corpo di Cristo.”
 (DP 1599: 110₁₃₋₁₄)
- (7) *Bet pařzink wiena alba du ghrieku*
 ma confessare-IMP 1-ACC oppure 2-ACC peccato-ACC.DU
 “Ma confessa uno o due peccati.”
 (VE 1579: 31₈)
- (8) [...] *turedams řchimta awiu /*
 avere-SMPRTC.NOM.SG.M cento-ACC pecora-GEN.PL.F
ir iey wiena pameřtū [...]
 e se 1-ACC.F perdere-SUBJ.PRS.3
 “[...] avendo cento pecore, se ne perdesse una [...]”
 (VEE 1579: L4v(88)₃₋₄; cfr. Lc 15,4)

La presenza del numerale cardinale ‘uno’ è regolare nelle costruzioni partitive:

- (9) *Bau kalbeiau efř kadařgi wienam*
 INT parlare-PST.1.SG io una volta 1-DAT.M
Sudzių Ifraelo
 giudice-GEN.PL.M Israele-GEN.SG
 “Ho io forse mai parlato a uno dei giudici di Israele.”
 (BrB 1590: 26v₇₋₉; cfr. 1Cr 17,6)
 cfr. ted. *Hab ich auch zu der Richter einem in Jsrael je gesagt*
 (Luth. 1545)
 lat. *numquid locutus sum uni iudicum Israel* (Vulg.)

3.2. Funzione nominale

Il passo (10) testimonia uno dei rari casi di funzione nominale del tipo *l'autobus (numero) cinque*. Qui il numerale ‘uno’ (così come ‘due’ e ‘quattro’) identifica (e non quantifica) il N cui si riferisce:

- (10) *tur* *gedoti* *Pfalma* *wiena* *alba*
 dovere-PRS.3 cantare-INF salmo-ACC.SG 1-ACC oppure
du *alba* *keturis* *pagalei* *priewoles*
 2-ACC oppure 4-ACC secondo necessità-GEN.SG
 “Deve cantare il Salmo uno, due o quattro secondo necessità.”
 (MžG II 1570: 495₇₋₈)

La rarità di casi come (10) nei testi cinquecenteschi non è dovuta a ragioni sintattiche: usi del genere sono tipici della funzione metatestuale, p. es. quando l'autore fa riferimento a un capitolo del libro o, come nel caso appena visto, a dei singoli componimenti (numerati). È chiaro che casi del genere sono rari in testi come questi, di argomento religioso, in cui la voce dell'autore-traduttore è ridotta al minimo e spesso confinata allo spazio paratestuale (dediche, prefazioni).

3.3. Funzione pronominale

In lituano antico il lessema ‘uno’ funziona anche come pronome correlativo e indefinito. Nella sua funzione pronominale, il lessema sostituisce un N che può essere o meno menzionato precedentemente nel testo.

3.3.1. Pronome correlativo

La funzione correlativa può manifestarsi nella relazione di opposizione (11-13) oppure di reciprocità (14-15).

3.3.1.a Opposizione

Nella correlazione di tipo oppositivo, ovvero quella in cui l’‘uno’ si oppone all’‘altro’ (o ‘secondo’), i due pronomi sono tipicamente anaforici e si riferiscono a degli elementi precedentemente introdotti nel discorso:

- (11) *bua* *du* *wiru,* *wienas* [...] *Antras*
 essere-PST.3 2-NOM uomo-DU 1-NOM.M
nekrikštitas, *Antras*
 NEG-battezzare-PTCP.PST.PASS.NOM.SG.M 2-ORD.NOM.M
krikštitas
 battezzare-PTCP.PST.PASS.NOM.SG.M
 “C’erano due uomini, uno [...] non battezzato, l’altro (letteralm. “il secondo”) battezzato.”
 (WP 1573: 163r₁₂₋₁₃)

- (12) *Jr ghis nutwere abu [...] Stulpu [...]*
 e egli-NOM afferrare-PST.3 entrambi-ACC.M colonna-ACC.DU.M
wieną ing sawa deschinę ranką, bei
 1-ACC.SG.M in proprio destro-ACC.SG.F mano-ACC.SG.F e
antrą ing sawa kairę ranką
 2-ORD.ACC.M in proprio sinistro-ACC.SG.F mano-ACC.SG.F
 “Ed egli afferrò entrambe [...] le colonne [...], una nella sua mano de-
 stra, e l'altra (letteralm. “la seconda”) nella sua mano sinistra.”
 (BrB 1590; Palionis e Žukauskaitė, 1983: 109; cfr. Gdc 16,29)

Tuttavia si danno anche casi in cui il pronome correlativo non è anaforico e gli elementi cui esso si riferisce sono desumibili solo dal contesto, oppure rimangono indefiniti:

- (13) *Wiens taip ir kits kitaip saka*
 1-NOM.M così e altro-NOM.SG.M diversamente dire-PRS.3
 “Uno dice in un modo e un altro in un altro.” (MžG II 1570: 349₈₋₉)

3.3.1.b Reciprocità

Il pronome correlativo è invece tipicamente non anaforico quando compare nelle costruzioni reciproche:

- (14) *prifakau iumus, idant wienas*
 comandare-PRS.1.SG voi-DAT affinché 1-NOM.M
antrą miletumbit
 2-ORD.ACC.SG.M amare-SUBJ.2.PL
 “Vi comando che vi amiati l'un l'altro (letteralm. “uno_[NOM] (il) secondo_[ACC]”).” (BrB 1590: 210v; cfr. Gv 15,17)
- (15) *ir ius wienas antro kaias*
 anche voi-NOM 1-NOM.M 2-ORD.GEN.SG.M piede-ACC.PL.F
turit mašgoti
 dovere-PRS.2.PL lavare-INF
 “Anche voi dovete lavare i piedi l'uno dell'altro.”
 (BrB 1590: 208v; cfr. Gv 13,14)

Tuttavia casi come (14-15) sono piuttosto periferici. Per l'espressione della reciprocità in lituano antico si ricorre più frequentemente

3.3.2. *Pronome indefinito*

(16) *Eikite* *Mieftana* *wienop*
andare-IMP.2.PL città-ILL.SG.M 1-ALL
“Andate in città da uno (un tale).” (BrB 1590: 60v; cfr. Mt 26,18)
cfr. ted. *Gebet hin in die Stad zu einem* (Luth. 1545)
lat. *Ite in civitatem ad quendam* (Vulg.)

- [illegible]

Casi come questi rientrano nella tipologia di uso indefinito-referenziale (o indefinito-specifico, ingl. *indefinite specific*), ovvero una situazione nella quale il referente è noto all'emittente ma non lo è al destinatario del messaggio (cfr. Holvoet e Tamulionienė, 2006: 18). Va sottolineato che questo impiego di *vienas* è poco frequente sia nei testi antichi che nel lituano moderno, dove più spesso il referente indefinito viene espresso attraverso un sostantivo

(p. es. *žmogus* “persona”) eventualmente (ma non obbligatoriamente) accompagnato da un aggettivo indefinito come *toks* “tale, simile”, *kažkoks* “certo, qualche” o – più raramente – proprio da *vienas* (v. sotto, § 3.5). Ad esempio il passo evangelico (16) nella traduzione moderna recita: *eikite į miestą pas tokį žmogų* “andate in città da una certa persona”; mentre il passo (17) recita: *mes matėme vieną tokį, kuris [...] išvarinėja demonus* “abbiamo visto un tale [...] che scaccia i demoni”.

Nei casi visti sopra l’uso di *vienas* sembra ricalcare quello del ted. *ein* e la ragione del suo impiego, soprattutto nella traduzione della Bibbia di Jonas Bretkūnas (Bretke, 1536-1602), è da ricercare proprio nell’influenza sintattica esercitata dal modello alto-tedesco: è noto, infatti, che il primo traduttore della Bibbia in lituano consultò fonti ebraiche, greche e latine, ma si basò soprattutto sulla traduzione di Lutero, come ben si nota confrontando il testo lituano con le corrispondenti precedenti versioni tedesca e latina.

Spesso, poi, *vienas* si va a comporre con le particelle *ne*, *al* e *kiek* per formare i pronomi e gli aggettivi indefiniti “nessuno”, “ciascuno”, “ognuno” secondo un procedimento che trova paralleli assai simili in latino (cfr. *ne ipse unus* e *quisque unus*) e in molte altre lingue (cfr. Haspelmath, 1997: 183-184), p. es. *negaletu ne wenas ius kaltinti* “nessuno potrebbe accusarvi” (MžK 1547: 30₈), *Diéwas kiekwiénam ifâke* (cfr. pol. ant. *Pan Bog każdemu rofkazal*) “Dio a ciascuno ha comandato” (DP 1599: 282₄₅), *Alwiens kerschts tur palaubima* (cfr. *all viede hat nu ein ende*) “ogni vendetta ha una fine” (MžG I 2164)¹¹.

3.4. Funzione aggettivale

Dopo i numerali e i pronomi, la terza classe lessicale che può contare la presenza di *vienas* è quella degli aggettivi. All’interno degli usi

¹¹ L’ultimo esempio è tratto da ALEW II: 1233. Si rimanda a questo lavoro (pp. 1233-1236) anche per le altre forme derivate dalla radice *vien-* attestate in lituano antico.

aggettivali si lasciano individuare tre sotto-tipi: (a) *descrittivo*, (b) *restrittivo*, (c) *dimostrativo*. Quella tra (a) e (b) è una distinzione su base semantica che generalmente non trova spazio nei lessici¹².

3.4.1. Funzione descrittiva

In questo impiego, il significato del lessema corrisponde a quello di "solo, da solo, privo di compagnia":

- (18) *neger jra ųmogui wienam buti*
 NEG-bene essere-PRS.3 persona-DAT.SG.M 1-DAT.M essere-INF
 "Non è bene per l'uomo essere uno (= solo)."

(VE 1579: 50₁₄₋₁₅; cfr. Gen 2,18)

cfr. ted. *Es ist nicht gut das der Mensch allein sey* (Luth. 1545)

lat. *Non est bonum esse hominem solum* (Vulg.)

- (19) *Er wiens walgijau ųawa kanųni,*
 INT 1-NOM.M mangiare-PST.1.SG proprio boccone-ACC.SG.M
er ne walge to ir
 INT NEG mangiare-PST.3 DEM.GEN.SG.M anche
Sirata?

orfano-NOM.SG

"Ho mangiato da solo il mio boccone, non ne ha mangiato anche l'orfano?"
 (BrB 1590: 171r₁₈₋₁₉; cfr. Gb 31,17)

cfr. ted. *Hab ich meinen bisßen allein gessen vnd nicht der Waise auch dauon gessen?* (Luth. 1545)

lat. *Si comedi buccellam meam solus, et non comedit pupillus ex ea* (Vulg.)

- (20) *nueya patųai wienas ant kalna*
 andare-PST.3 DEM.NOM.SG.M 1-NOM.M su monte-GEN.SG
 "Andò da solo sul monte." (VEE 1579: G2r₁₆; cfr. Gv 6,15)

cfr. ted. *entweich er abermal auff den Berg er selbs alleine* (Luth. 1545)

lat. *secessit iterum in montem ipse solus* (Vulg.)

¹² I tre sotto-tipi descrittivo, restrittivo e dimostrativo corrispondono, nella terminologia di KUTEVA *et al.* (2019: 298-306), ai processi di grammaticalizzazione: *one (numeral) > alone; one (numeral) > only; one (numeral) > same*.

Questi passi – e altri in cui si ravvisa questa stessa funzione – se confrontati con le fonti, mostrano una corrispondenza regolare tra il lit. *vienas*, il ted. *allein* e il lat. *sōlus*. Il tedesco e il latino infatti non prevedono l'impiego aggettivale di *ein* e *ūnus* con questo valore.

3.4.2. Funzione restrittiva

Impiegato in funzione restrittiva, 'uno' isola l'elemento a cui si riferisce come l'unico soggetto o oggetto dell'azione espressa dal predicato. Nei testi antichi questa funzione è più frequente della precedente. La si incontra riferita tanto al soggetto (21-22), quanto al complemento diretto (23) e a quello indiretto (24-25):

- (21) *Tu atpraſchei mums wienas Sawa*
 tu pregare-PST.3 noi-DAT 1-NOM.M proprio-ACC.SG.M
Dangaus Tiewa
 cielo-GEN.SG.M padre-ACC.SG.M
 "Tu solo hai pregato per noi il tuo Padre del cielo."
 (MžG I 1566: 221₁₆₋₁₇)

- (22) *gifsai wenas tur buti*
 egli 1-NOM.M dovere-PRS.3 essere-INF
garbinamas [...] *a ne*
 glorificare-PTCP.PRS.PASS.NOM.SG.M ma NEG
kittas kas
 altro-NOM.SG.M qualcuno-NOM.SG.M
 "Egli solo deve essere glorificato [...] e non qualcun altro."
 (WP 1573: 89_v₁₃₋₁₄)

- (23) *turrimē ghi wenū ſchlawinti garbinti,*
 dovere-PRS.1.PL egli-ACC 1-ACC.M adorare-INF glorificare-INF
ghi wenū melſti
 egli-ACC 1-ACC.M pregare-INF
 "Dobbiamo adorare e glorificare lui solo, pregare lui solo."
 (WP 1573: 5r₁₆₋₁₇)

- (24) *Tikekem mes daugefnei wienai Mariey*
 credere-IMP.1.PL noi di più 1-DAT.F Maria-DAT
 "Crediamo di più alla sola Maria." (MžG II 1570: 255₁₃₋₁₄)
- (25) *Chriftus wiënam Pêtrui awis*
 Cristo-NOM 1-DAT.SG.M Pietro-DAT pecora-ACC.PL
sawás padūdamas
 proprio-ACC.PL dare-SMPRTC.NOM.1.SG
 "Cristo, affidando al solo Pietro le sue pecore." (DP 1599: 611₃₂₋₃₃)
 cfr. pol. ant. *Pan Chryſtus iednemu Piotrowi owce fwe porucżaiac*

In coda a questi esempi, si noti come la funzione restrittiva avvicini il tipo aggettivale (*solo, unico*) a quello avverbiale (*solamente, soltanto*). Il fenomeno è noto anche alla lingua latina, p. es. *unos ex omnibus Sequanos* "solo i Sequani fra tutti" (Caes. *De bello gallico* I.32); *unas mihi scito litteras redditas esse* "sappi che mi è stata consegnata solamente/una sola lettera" (Cic. *Att.* 1.5.4); *unae quinque [...] minae* "solamente cinque [...] mine" (Pl. *Ps.* I.I.52). Allo stesso tempo è anche evidente che in molte lingue la distinzione tanto semantica quanto formale fra l'aggettivo *solo* e l'avverbio *soltanto* è molto ridotta o perfino assente (p. es. lat. *sōlus* : *sōlum*; ted. *allein* : *allein/nur*). Le traduzioni lituane dell'aggettivo latino *sōlus* in funzione restrittiva sono rese ora con l'aggettivo *wienas*, ora con l'avverbio *tik/tiktai* "soltanto". Esempi della corrispondenza lat. *ūnus* : lit. ant. *wienas* sono il passo (24) che risponde al lat. *Credendum est magis soli Mariae*; oppure *Pona Dieua tawa garbiſi ir iem wienam tarnauſi* (VaišŽT 1600: 188b₁₇₋₁₈) "onorerei il Signore tuo Dio e lui solo servirai", cfr. lat. *DEUM tuum adorabis, et illi soli servies*. Invece esempi della corrispondenza lat. *ūnus* : lit. ant. *tik/tiktai* sono: *ne per Zokona tiektai* "non soltanto per la Legge", cfr. lat. *neque per Legem ſolam*; oppure *tiktai pawadintumpiump* (VaišŽT 1600: 76a₁₁) "soltanto ai chiamati (letteralm. "ai nominati")", cfr. lat. *ad solos electos*. Una testimonianza ancora più chiara di come le due costruzioni lituane siano intercambiabili è offerta da Mažvydas, il quale nel Catechismo scrive: *KRikſtas ne eſti tektai praſtas wānduo* (MžK 1547: 24₃) "il Battesimo non è solamente semplice acqua" e dodici anni più tardi, nella *Forma Chrikſtima*, scrive: *Bet ne eſt wiens wandū praſtas* (MžFK 1559: 13₁₉ = MžGK I 1566: 233₉)

“ma non è solamente acqua semplice”. Le due versioni mažvydiane, viste in parallelo, rivelano l’uso di *wiens* come sinonimo dell’avverbio *tektai* (lit. mod. *tiktai*) “solamente”¹³.

Dall’aggettivo *vienas* – più precisamente, dal nom-acc.sg. neutro *viena* (cfr. Smoczyński, 2018: 1651) – si è poi sviluppata, con apocope, la particella avverbiale *vien* “soltanto, solamente” che non solo è sinonimo dell’avverbio *tik/tiktai*, ma che spesso compare assieme ad esso nella locuzione avverbiale *vien tik* “solamente” (cfr. LKŽe, s.v.).

3.4.3. Funzione dimostrativa

Sebbene più raramente, *vienas* può essere utilizzato anche con il valore dimostrativo “stesso, medesimo”:

- (26) *wiena* *yrà* *wissu* *pabaga*
 1-NOM.F essere-PRS.3 tutto-GEN.PL fine-NOM.SG.F
 “Una (= la stessa) è la fine di tutti.” (DP 1599: 569₄₁)
 cfr. pol. ant. *iedno iest wśytkich dokończenie*

A differenza di quanto si nota per il polacco, dove *jeden* svolge questa stessa funzione, in (27-28) le fonti latine e tedesche mostrano l’impiego di lessemi diversi dal numerale ‘uno’; in particolare, il latino presenta l’aggettivo dimostrativo *idem*, mentre il tedesco l’aggettivo invariabile (non dimostrativo) *einerlei* che vale piuttosto ‘dello stesso genere, identico’ (sinonimo di *gleichartig*):

- (27) *Wienas* *testow* *Prisakimas*
 1-NOM.SG IMP-stare-PRS.3 comandamento-NOM.SG
 “Ci sia una sola (= la stessa) Legge.”
 (BrB 1590; Palionis e Žukauskaitė, 1983: 58; cfr. Es 12,49)
 cfr. ted. *Einerley Gesetz sey* (Luth. 1545)
 lat. *Eadem lex erit* (Vulg.)

¹³ Questo passo si presta alla comparazione diretta con il sintagma *ains tickars vnds* attestato nell’*Enchiridion* prussiano antico (Ench. 1561: 6120). I paralleli lituani offrono sostegno alla tesi di DINI (2014: 64) secondo cui il numerale prussiano *ains* “uno” vada inteso, in quest’unico caso, «im Sinne von “allein, nur eins”». Perciò tanto il pr. ant. *ains tickars vnds* quanto il lit. ant. *wiens wandũ praštas* si lasciano tradurre come “solamente semplice acqua”.

- (28) *a ne wiffi sanarei wiena*
 ma NEG tutto-NOM.PL membro-NOM.PL 1-ACC.M
darba tur
 lavoro-ACC.SG avere(-PRS.3)
 “Ma non tutte le membra hanno la medesima funzione.”
 (VEE 1579: C4r(23)₁₃₋₁₄; cfr. Rom 12,4)
 cfr. ted. *aber alle glieder nicht einerley Geschefft haben* (Luth. 1545)
 lat. *omnia autem membra non eundem actum habent* (Vulg.)

3.5. Funzione di pseudo-articolo

Di seguito si discutono alcuni casi in cui il primo numerale sembra prestarsi alla funzione di articolo indefinito. Se da un punto di vista teorico e tipologico è cosa del tutto comune che gli articoli indefiniti, nelle lingue che li possiedono, si sviluppino a partire dal numerale ‘uno’, ciò che lascia perplessi è che il lituano non ha (né ha mai avuto) gli articoli. L’uso di ‘uno’ in funzione di articolo indefinito del resto non è segnalato né dalle grammatiche antiche, né dagli studi moderni.

Il primo esempio è tratto dal secondo tomo delle *Gesmes chrikščioniškas* (“Inni cristiani”) di Mažvydas, una raccolta di preghiere, inni e canti religiosi a uso liturgico spesso tradotti dal latino, dal tedesco e dal polacco¹⁴:

- (29) *Buwa wienas wijras teifus*
 essere-PST.3 1-NOM.M uomo-NOM.SG.M giusto-NOM.SG.M
 “C’era un uomo giusto.” (MžG II 1570: 430₇)

Nello specifico, questo è l’inno a San Giobbe (talvolta citato con le parole iniziali *Wis galijfs Diewe tieve*), un canto in quartine che narra le vicende del personaggio biblico. Il passo (29) apre la seconda quartina, che recita così:

¹⁴ Per le fonti degli inni di Mažvydas cfr. MICHELINI (2000; 2002). I due tomi degli *Inni cristiani* di Mažvydas furono pubblicati postumi sotto la cura di Baltramiejus Vilentas (ca. 1525-1587), cfr. FORD (1969: 16).

- (30) *Buwa wienas wijras teifus*
Praštibiejė sawa labai baišus
kurio buwa wardas Jobas
*Po wiffa fchi fwieta ape fkielbtas*¹⁵

È fuor di dubbio che *wienas* si riferisca a *wijras* con cui è concordato, così come è evidente che la sua funzione non è aggettivale – vuoi descrittiva (quell'uomo giusto non era *da solo*), vuoi restrittiva (non era *l'unico*). La funzione di 'uno' non è neanche cardinale poiché esso non quantifica il referente, bensì ne presenta uno *ex novo*. Qui *wienas* è molto simile a un articolo indefinito: introduce un nuovo N di cui ancora non è stato detto niente e di cui si parlerà nel prosieguo.

L'esposizione di questi versi naturalmente si rifà all'*incipit* del libro biblico di Giobbe (Gb 1,1), di cui si riporta la versione greca dei Settanta (31a), quella latina della *Vulgata* (31b) e quella della Bibbia di Lutero (31c):

- (31a) Ἀνθρωπὸς τις ἦν ἐν χώρᾳ τῇ Αὐσίτιδι, ὃ ὄνομα Ἰωβ· καὶ ἦν ὁ ἄνθρωπος ἐκεῖνος ἀληθινός, ἀμεμπτος, δίκαιος, θεοσεβής, ἀπεχόμενος ἀπὸ παντὸς πονηροῦ πράγματος. (LXX)
- (31b) *Vir erat in terra Us nomine Iob, et erat vir ille simplex et rectus ac timens Deum et recedens a malo.* (Vulg.)
- (31c) *ES war ein Man im lande Vz der hies Hiob Derselb war schlecht vnd recht Gottfürchtig vnd meidet das böse.* (Luth. 1545)

Il testo greco ricorre a *τις* per marcare l'indefinitezza del N. In latino, lingua senza articoli, *vir* è introdotto senza ulteriori elementi (qui sarebbe stato possibile, secondo il modello dei LXX, avere *quīdam*), mentre in tedesco *Man* è naturalmente preceduto dall'articolo indefinito. A rendere ancora più insolita l'attestazione di questo 'pseudo-articolo' nell'inno a San Giobbe c'è il fatto che, se si va a cercare il passo biblico (Gb 1,1) nella coeva *Postilla* anonima detta 'di Wolfenbüttel' non c'è traccia di *vienas*:

¹⁵ In traduzione italiana: "C'era un uomo giusto / nella sua semplicità assai timorato / il cui nome era Giobbe / rinomato in tutto questo mondo".

- (32) *Bua* *wiras* *ŝzemeia* *Hus*
 essere-PST.3 uomo-NOM.SG.M terra-LOC.SG.F Uz
wardu *Job*
 nome-INS.SG.M Giobbe
 "C'era (un) uomo nella terra di Uz di nome Giobbe."
 (WP 1573: 292_r¹⁹⁻²⁰)

Lo stesso vale per le traduzioni lituane della Bibbia prodotte nei secoli XVI-XVIII, quali quella di Bretkūnas (1590) (33a), di Chylinksis (1660) (33b) e di Šimelpenigis (1755) (33c):

- (33a) *Buwa Wiras Szemeie Vz, wadinamas Hiobas* (BrB 1590: 48r)
 (33b) *Buwo źmogus źiamey Uz, wardas jo buwo Job* (ChB 1660: 388)
 (33c) *Buwo wyras źemeje Uz, wardu Jobas* (ŠB 1755: 712)

Si tenga anche conto che costrutti simili, in altri passaggi biblici, figurano senza pseudo-articolo:

- (34a) *O buwa Wiras*
 "E c'era un uomo." (BrB 1590; Palionis e Žukauskaitė, 1983: 99)
 (34b) *Bei Nueia Wiras isch Namu Leui, ir waede Dukterį Leui*
 "Un uomo della casa di Levi andò e sposò una figlia di Levi."
 (BrB 1590; Palionis e Žukauskaitė, 1983: 36)

Per spiegare l'insolita presenza di *wienas* nell'inno lituano antico a San Giobbe occorre innanzitutto rintracciarne le fonti. Secondo Michellini (2002: 67) questo inno «può venir considerato una libera rielaborazione in chiave luterana del libro di Giobbe, pur essendo innegabile che la sua parte iniziale si ricollega al componimento polacco *Wszechmogący Pánie miły*, in diverse frasi tradotto senza grossi cambiamenti». Eccone il testo riportato da Michellini (2000: 492):

- (35) *Był mąż ieden sprawiedliwy,*
W prostoci swej bázro boiazliwy,
Ktoemu Jop imię było,
Po wszystkim świecie o nim słynęło.

La versione polacca mostra proprio la presenza del numerale *jeden* “uno” posposto al N *maq̣* “uomo”. Va notato che tale costrutto è a sua volta marcato, o quantomeno divergente rispetto alle traduzioni polacche della Bibbia sia antiche che moderne. Tra quelle in circolazione all’epoca, di particolare rilievo fu la Bibbia cosiddetta “di Brest” (1563). Eccone l’*incipit* del libro di Giobbe:

- (36) *Byl w̋ziemi Uz Ma̋ nie khtory imieniem Job*
 “C’era nella terra di Uz un certo (*nie khtory*) uomo (*Ma̋*) di nome
 Giobbe.” (BB 1563: 274r)

Qui non c’è *jeden*, mentre l’indefinitezza del N è marcata dall’aggettivo indefinito posposto *nie khtory* (in maniera molto simile alla versione dei Settanta: ἄνθρωπός τις).

Mettendo insieme i dati raccolti, emerge un quadro bipartito in due filoni compatti (v. Tabelle 2 e 3): da un lato la tradizione biblica, dall’altro quella innologica. Nel primo caso le attestazioni rispecchiano ciò che sappiamo sui sistemi linguistici di riferimento: le lingue senza articolo non marcano espressamente l’indefinitezza del N (o il polacco lo fa con un aggettivo indefinito, come il greco) mentre il tedesco utilizza l’articolo *ein*. Nel secondo caso invece incontriamo un inconsueto pseudo-articolo la cui anomalia è sottolineata dalla sua assenza nei coevi testi biblici.

LINGUA	Fonte	TESTO
gr.	LXX	Ἄνθρωπός τις ἦν ἐν χώρᾳ τῇ Αὐσίτιδι
lat.	Vulg.	<i>Vlrat in Terra Hus</i>
ted.	Luth. 1545	<i>ES war ein Man im lande Vz</i>
pol. ant.	BB 1563	<i>Byl w̋ziemi Uz Ma̋ nie khtory</i>
lit. ant.	BrB 1590	<i>Buwa Wiras Szemeie Vz</i>
	WP 1573	<i>Bua wiras ſzemeia Hus</i>
	ChB 1660	<i>Buwo źmogus źiamey Uz</i>
	ŠB 1755	<i>Buwo wyras źemeje Uz</i>

Tabella 1. *Gb 1,1: tradizione biblica.*

LINGUA	Fonte	TESTO
pol. ant.	<i>Wszzechmogący Pánie miły</i>	<i>Był mąż ieden sprawiedliwy</i>
lit. ant.	MžGK II 1570	<i>Buwa wienas wijras teifus</i>

Tabella 2. *Gb 1,1: tradizione innologica.*

Vista la stretta somiglianza generale fra il testo lituano e quello polacco, il fenomeno dello pseudo-articolo nel primo si può ragionevolmente spiegare come un'interferenza da parte del secondo. Questo ci conduce all'origine del fenomeno (anche se si potrebbe ulteriormente chiedere ragione dell'introduzione di *jeden* nell'inno polacco: forse a sua volta interferenza da parte di una precedente versione tedesca con *ein*?), ma non esaurisce il problema. Infatti, rimanendo sul lituano, e pur ammettendo un'interferenza interlinguistica, restano alcune domande aperte:

- Perché lo stesso fenomeno non si incontra nelle traduzioni bibliche coeve o successive?
- Perché Mažvydas avrebbe qui tradotto pedissequamente? È noto che egli fu un traduttore tutt'altro che passivo; per di più, proprio in questo stesso inno egli aggiunse autonomamente alcune strofe che non trovano corrispettivo né nel libro di Giobbe né nella fonte polacca (Michelini, 2000: 492). Questo lascia aperta la ricerca di altre possibili fonti; in ogni caso, un atteggiamento acritico da parte di Mažvydas lascia senz'altro perplessi.
- Se si tratta di un semplice calco sul polacco, come spiegare la diversa posizione che *wienas* e *ieden* occupano rispetto a *wijras* e *mąž*? Lo spostamento dello pseudo-articolo prima del N è segnale di una certa integrazione di questo elemento nel sistema linguistico di arrivo.

L'individuazione della fonte polacca antica – non certa, ancorché probabile – non offre dunque ragione dell'integrazione del fenomeno. Per questo è opportuno seguire anche un'altra direzione, ovvero contestualizzare il passo nella tipologia testuale cui esso appartiene.

L'inno a San Giobbe – e in generale gli *Inni cristiani* – appartiene a un genere con delle caratteristiche ben precise. Le più rilevanti sono le seguenti:

- *Tono popolare*. Gli inni, soprattutto se dedicati a santi o personaggi biblici, lodano il protagonista proponendolo a modello di edificazione per i fedeli. Spesso rielaborano in forma divulgativa una o più vicende bibliche¹⁶.
- *Indipendenza*. Gli inni mostrano un certo grado di libertà nella gestione della materia. Non essendo traduzioni, ma piuttosto trasposizioni, essi procedono in modo alquanto indipendente dai testi canonici, i quali fanno da riferimento generale ma non vengono osservati alla lettera. Ad esempio, il terzo verso della quartina sopra citata (*Po wiffa ſchi ſwieta ape ſkielbtas* “rinomato in tutto questo mondo”) non trova corrispondenza nel testo biblico (ma Mažvydas evidentemente l'ha ripreso dal polacco *Po wszystkim świecie o nim słynęło*).
- *Fruizione orale*. Come indicano i frontespizi delle raccolte¹⁷, la destinazione principale di questi testi era l'esecuzione cantata e collettiva nelle celebrazioni liturgiche.
- *Struttura poetica*. I versi sono in rima o almeno assonanti, inoltre il testo deve concordare con la melodia, pertanto i singoli versi rispondono a una determinata griglia metrica.

Sono tutti aspetti che fanno dell'inno una tipologia testuale assai diversa da quella biblica. È evidente che occorre uscire dal circuito comparativo fra testi letterari ‘alti’ quali sono le traduzioni della Bibbia e guardare ai generi più vicini all'oralità quali, appunto, gli inni e i canti devozionali, senza peraltro trascurare quel vasto serbatoio di cultura tradizionale che sono le favole e i racconti popolari. Com'è

¹⁶ La tradizione cristiana (sia cattolica che protestante) è ricca di esempi simili. Passando all'italiano, ma restando su Giobbe, si possono richiamare diversi volgarizzamenti, ovvero rielaborazioni poetiche del testo biblico, quali il *Volgarizzamento in terza rima del sacro libro di Giob...* di Marcantonio Talleoni (Tal. 1764) o *Il libro di Giobbe volgarizzato per Giuseppe Costa* (Cos. 1848).

¹⁷ Ad esempio il frontespizio del primo tomo degli *Inni cristiani* di Mažvydas (1566) recita: “Inni cristiani cantati nelle chiese durante l'Avvento e il Natale fino alla Candelora”.

noto, questi generi si sono tramandati lungo i secoli in forma orale per trovare una fissazione scritta soltanto in epoca moderna. Per ciò che interessa qui, è opportuno notare che gli incipit delle fiabe tradizionali lituane – le *pasakos* – mostrano spesso uno pseudo-articolo in funzione presentativa (37a-f)¹⁸:

- (37a) *Vienas vargingas žmogus turėjo vienintelį sūnų* [...]

“Un pover'uomo aveva un unico figlio [...]"
- (37b) *Vienas neregys tėvas išsiuntė savo tris sūnus* [...]

“Un padre cieco inviò i suoi tre figli [...]"
- (37c) *Vienas bernas ilgą laiką tarnavo pas vieną ūkininką* [...]

“Un bracciante per lungo tempo servì presso un padrone [...]"
- (37d) *Buvo vienas turtingas ponas.*

“C'era un ricco signore.”
- (37e) *Gyveno vienas karalius.*

“Viveva un re.”
- (37f) *Vienas karalius gyveno ir turėjo sūnų.*

“Viveva un re e aveva un figlio.”

Alla luce di tutto ciò, non è forse un caso che lo pseudo-articolo *mažvydiano* sia testimoniato proprio in un testo come l'inno a San Giobbe e non nelle traduzioni bibliche. L'inno condivide con il genere dei racconti tradizionali almeno tre dei quattro aspetti sopra elencati: esso appartiene a un genere popolare, ammette una discreta libertà creativa ed è destinato all'esecuzione orale. Anche la quarta caratteristica, ovvero la struttura poetica, può aver giocato un ruolo nell'inserimento dello pseudo-articolo come 'zeppa' allo scopo di allungare il versetto. Infatti i primi versi di ogni strofa contano otto sillabe (così come quelli polacchi):

- (38) *Wis galijfis Dieve tieve* (strofa 1)
- Buwa wienas wijras teifus* (strofa 2)
- Tassai ing Diwa tikiėja* (strofa 3)
- Turieija funus septinis* (strofa 4)
- Sheimina beškaicziaus buwa* (strofa 5)
- Po Saules ufztekieghima* (strofa 6)
- ecc.

¹⁸ Esempi tratti da ALEKSYNAS (1985, a cura di).

Il versetto in questione, privato dello pseudo-articolo, risulterebbe ipometro. È possibile dunque che anche le necessità metriche abbiano spinto verso l'inserimento di un elemento di per sé non necessario ma tuttavia ammesso, specialmente nei registri non alti e nei contesti d'uso più orientati al parlato (o al cantato).

L'inquadramento filologico-testuale del passo (29) può ora essere utilmente integrato da considerazioni di ordine più strettamente linguistico. Come illustrato per primo da Givón (1981), nelle lingue del mondo il passaggio di 'uno' da numerale ad articolo indefinito si svolge lungo un *continuum*, o scala graduale, che prevede tre stadi corrispondenti alle seguenti funzioni: (i) quantificazione; (ii) referenzialità/denotazione; (iii) genericità/connotazione. Le idee di Givón sono poi state riprese e sviluppate in particolare da Heine (1997) e Schroeder (2006), i quali hanno proposto dei propri modelli più dettagliati¹⁹. Heine (1997: 66-82), nello specifico, delinea un percorso in cinque fasi: (i) *Numeral*; (ii) *Presentative marker*; (iii) *Specific marker*; (iv) *Non-specific marker*; (v) *Generalized article*. Come ricorda Napoli (2013: 191): «the use of 'one' as a *presentative marker* is proper to languages which do not have a grammaticalized indefinite article, and at an early phase is typically confined to the beginning of a tale».

Appare dunque corretto collocare l'*incipit* dell'inno di San Giobbe entro questa cornice e interpretare la funzione dello pseudo-articolo *vienas* come testimonianza di un processo incipiente di grammaticalizzazione del numerale 'uno': visto sotto questa luce, (29) rappresenta un'istanza del secondo stadio del modello di Heine. Il contesto tipico di questa funzione è quello delle proposizioni esistenziali / presentative. Secondo la definizione di McNally (2019: 281), tali costrutti "esprimono una proposizione circa l'esistenza o la presenza di qualcuno o qualcosa".

Questo tipo di predicazione non verbale, talvolta chiamata *inverse-locational*, si lascia ricondurre al modello *There is N (Loc)*, dove Loc rappresenta «a phrase expressing location» (Creissels, 2019). All'interno di questa cornice rientrerebbero anche le costruzioni esistenziali

¹⁹ Per un'esposizione critica dei tre modelli teorici e dei loro rapporti cfr. NAPOLI (2013: 189-192).

del tipo *There is N* – dove l’esistenza è interpretata come «being an element of the world» (cfr. Creissels, Bertinetto e Ciucci, *in stampa*)²⁰. Ora, la caratteristica dei costrutti esistenziali / presentativi è quella di essere *sentence-focus*, ovvero di mostrare una struttura informativa focalizzata sia sul predicato che sul soggetto (cfr. Lambrecht, 2000). In altre parole, essi svolgono la funzione di introdurre un nuovo referente focale o una nuova situazione all’interno del discorso; per questa ragione i costrutti presentativi pongono restrizioni interlinguisticamente variabili sulla compatibilità con SN definiti.

Già nel lituano del Cinquecento il lessema ‘uno’ può essere utilizzato in funzione di *presentative marker* per introdurre nel discorso un referente nuovo (indefinito). Questa, va ribadito, è solo un’opzione, infatti in molti casi la funzione presentativa non ha marche morfologiche esplicite e consiste nel solo verbo essere in funzione esistenziale anteposto al SN (v. Tabella 1)²¹.

I seguenti esempi (39-41) testimoniano un passaggio ulteriore, ovvero l’utilizzo di ‘uno’ in funzione di *specific marker*. In questa funzione, esso viene tipicamente applicato a quegli elementi del discorso che sono noti all’emittente del messaggio ma che si presume non siano noti al destinatario (cfr. Heine, 1997: 72-73). È questo un uso referenziale (v. il secondo stadio di Givón, 1981) dal momento che l’emittente si riferisce a un elemento specifico (e non ipotetico) a lui noto:

- (39) *yra cže wiėnas iaunkáitis*
 essere-PRS.3 qua 1-NOM.M giovinetto-NOM.SG.M
 “C’è qua un giovinetto.” (DP 1599: 125₃; cfr. Gv 6,9)
 cfr. pol. ant. *Bà iefłci tu iedeń młodzienc*

²⁰ Ringrazio il dott. Luca Ciucci per aver messo a mia disposizione un’anteprima di questo lavoro.

²¹ Stando alla rassegna tipologica di CREISSELS (2019), oltre la metà delle lingue del mondo non avrebbe marche dedicate per la predicazione esistenziale (*inverse locational*). In queste lingue la prospettiva esistenziale (*There is N Loc*) viene distinta da quella locativa (*N is Loc*) attraverso un diverso ordine dei costituenti. Il lituano rientra in questo tipo, infatti nella predicazione esistenziale il N (indefinito) segue il verbo: (a) *gyveno vyras Uco šalyje* “viveva (c’era) un uomo nel paese di Uz”; al contrario, nella predicazione locativa il N (che può essere definito o indefinito) precede il verbo: (b) *vyras gyveno Uco šalyje* “un/l’uomo viveva nel paese di Uz”.

- (40) *O ųtoios wieną diena [...]*
 ma stare-PST.REFL.3 1-ACC?.F giorno-NOM.SG.F
 “E un (certo) giorno [...]” (BrB: 148r₂₆; cfr. Gb 1,6)
 cfr. ted. *ES begab sich aber auff einen tag* (Luth. 1545)
 lat. *Quadam autem die* (Vulg.)
- (41) *Ir kaip ghis ateia y wiena*
 e come egli arrivare-PST.3 in 1-ACC.M
mieųteli
 città-DIM.ACC.SG.M
 “E come egli arrivò in un (certo) villaggio.”
 (VEE 1579 O3r(109)5; cfr. Lc 17,12)
 cfr. ted. *Vnd als er in einen Marckt kam* (Luth. 1545)
 lat. *Et cum ingrederetur quoddam castellum* (Vulg.)

Visti in parallelo con le loro fonti, i passaggi lituani sembrano ricalcare le strutture del polacco e del tedesco. Tuttavia la presenza di *vienas* non si lascia spiegare soltanto come risultato di interferenza traduttiva. Se da un lato è certamente possibile che i traduttori siano stati spinti all’inserimento di *vienas* dai testi modello cui facevano riferimento, dall’altro è fondato ritenere che l’uso di ‘uno’ in funzione indefinita referenziale facesse parte del sistema grammaticale del lituano nella seconda metà del XVI secolo. A far propendere verso questa lettura sono due considerazioni. La prima è che nella maggior parte dei casi analoghi, anche in corrispondenza di *jeden* e di *ein*, il lituano risponde con Ø; da ciò deriva che l’influenza sintattica delle lingue modello – almeno limitatamente al fenomeno qui indagato – non può dirsi particolarmente forte (a differenza di quanto è dato osservare nei coevi testi in prussiano antico e in lettone antico)²². In

²² Per quanto riguarda il prussiano antico, alcuni autori sostengono che il numerale *ains* “uno” nel XVI secolo era ormai già passato alla funzione di articolo indefinito. Sebbene non originario, in questa lingua l’articolo avrebbe finito per stabilizzarsi e funzionare regolarmente, cioè in modo indipendente dal tedesco. La questione è riassunta in PARENTI (1995), il quale porta argomenti a sostegno dell’ipotesi che gli articoli definiti e indefiniti facessero parte del sistema linguistico del prussiano antico. Anche i testi lettoni del Cinquecento e del Seicento, che furono per lo più traduzioni dal tedesco operate da traduttori di lingua madre tedesca, mostrano una dipendenza molto forte dalle strutture (morfo-)sintattiche del testo modello. Qui si osserva una corrispondenza

secondo luogo, l'uso di 'uno' come *specific marker* è attestato lungo tutta la storia della lingua lituana (cfr. LKŽe, *s.v. vienas* 2.6) ed è tutt'ora vivo, sebbene in questa funzione *vienas* alterni con Ø e con gli aggettivi *kažkoks* e *toks* (cfr. Holvoet e Tamulionienė, 2006: 18; Spraunienė, 2008).

4. Conclusioni

L'analisi fin qui condotta ha permesso di descrivere la varietà di funzioni che caratterizza il lessema 'uno' nei testi lituani del XVI secolo, ovvero durante la prima stagione della scrittura in questa lingua. Oltre alla funzione primaria e prototipica, cioè quella di numerale cardinale, il quadro complessivo degli utilizzi di *vienas* comprende altre quattro funzioni non-numerali: quella nominale, quella di pronome, di aggettivo e di pseudo-articolo – le ultime tre ulteriormente articolate in sotto-funzioni; cfr. Tabella 3).

La prima funzione, di cui il corpus analizzato offre pochi esempi, consiste nell'assegnazione di un'etichetta nominale, per cui il lessema passa da numerale a nome proprio.

Usato come pronome, *vienas* può essere sia correlativo che indefinito. Nella funzione correlativa si trova tanto in costruzioni oppositive che reciproche. Nelle costruzioni oppositive il pronome è tipicamente (ma non obbligatoriamente) anaforico, mentre in quelle reciproche esso è tipicamente non anaforico. L'impiego di *vienas* è più frequente nelle costruzioni oppositive, mentre in quelle reciproche è più comune il ricorso a sintagmi del tipo preposizione + pronome (p. es. “tra sé, tra loro”), oppure l'impiego del pronome *kitas* “altro”. Questo, che tipicamente funge da secondo membro della correlazione, va a ricoprire anche la posizione di primo membro (“l'altro [...] l'altro”).

molto regolare tra l'articolo indefinito tedesco e il numerale *wens* (lett. mod. *viens*). Tra gli esempi presentati in CERRI (2016: 221-222) ve ne sono alcuni che rientrano nella funzione indefinita non-referenziale (*non-specific marker*) e perfino in quella di articolo indefinito vero e proprio (*generalized article*), ovvero le tipologie (iv) e (v) dello schema di HEINE (1997: 72-73).

La funzione di pronome indefinito è infrequente e limitata al riferimento indefinito specifico (referenziale). Nei testi cinquecenteschi tradotti dal tedesco usi di questo tipo sembrano essere indotti – o almeno favoriti – dalla presenza di *ein* nel testo modello. Per comprendere meglio il ruolo dell'interferenza linguistica è auspicabile in futuro un'analisi comparata tra i testi tradotti dal tedesco e quelli tradotti dal polacco.

Passando alla funzione aggettivale, *vienas* può assumere tre diversi valori. Il significato descrittivo vale "solo, privo di compagnia"; quello restrittivo invece vale "unico" ed è attestato più frequentemente del precedente. Nei testi tradotti da fonti latine e/o tedesche si osserva che *vienas* corrisponde al lat. *sōlus* e al ted. *allein*. In questa funzione restrittiva, l'aggettivo *vienas* alterna sinonimicamente con l'avverbio *tik/tiktai* "soltanto" (alternanza, questa, documentata anche in lettone antico e, in un caso, anche in prussiano antico). Il terzo valore della funzione aggettivale è quello dimostrativo "stesso, medesimo". Anche questo impiego di *vienas* non dipende dall'influenza delle lingue modello, le quali in questa funzione utilizzano altri mezzi (lat. *īdem*, ted. *einerlei*).

L'ultima funzione documentata è quella di pseudo-articolo, ovvero di marca di indefinitezza del sintagma nominale. Il lituano, lingua priva di articoli, nella maggior parte dei casi non marca esplicitamente l'indefinitezza. Per spiegare i casi in cui *vienas* funziona come marca esplicita si può ricorrere a due tipi di considerazioni. Il primo tipo, di ordine filologico-testuale, si rivolge alle fonti e alle caratteristiche stilistico-metriche del testo. Si è visto che *vienas* risponde al pol. *jeden* e al ted. *ein*, perciò il suo inserimento può essere stato favorito dal testo modello. In certi casi anche la tipologia testuale può avere agevolato l'inserimento di un elemento non obbligatorio ma ammesso soprattutto nei registri più popolari e vicini al parlato; infatti se ne riscontrano paralleli lungo la tradizione secolare del racconto e della fiaba. Qui intervengono le considerazioni del secondo tipo, ovvero tipologiche. Compreso in questo contesto, 'uno' funziona come la marca presentativa tipica dei costrutti esistenziali / presentativi frequenti negli *incipit* dei racconti.

CLASSE LESSICALE	FUNZIONE NUMERALE	FUNZIONI NON-NUMERALI			
		Nome	Pronome	Aggettivo	Pseudo-articolo
FUNZIONE (SOTTO-FUNZIONE)	Numeraledtd>	Nominale (identifica)	Correlativa oppositiva (“l’uno... l’altro”)	Descrittiva (SN è “da solo”)	Presentativa (introduce un SN nuovo e indefinito)
	–	–	Correlativa reciproca (“l’un l’altro”)	Restrittiva (SN è “l’unico”)	Indefinita specifica (“un certo” SN noto all’emittente)
	–	–	Indefinita (“un certo”)	Dimostrativa (SN è “uno stesso”)	–

Tabella 3. Sinossi: polifunzionalità di ‘uno’ in lituano antico.

In altri casi ancora, 'uno' funziona come marca di indefinitezza specifica (referenziale) del sintagma nominale. È un impiego poco frequente; molto più spesso questo tratto non è segnalato da marche esplicite.

Dal punto di vista tipologico, 'uno' in lituano antico presenta un grado di grammaticalizzazione limitato. Facendo riferimento al modello a cinque stadi di Heine (1997), *vienas* può fungere da marca presentativa (stadio ii), e può accompagnare i referenti indefiniti specifici (stadio iii) – sempre però in alternativa alla ben più frequente assenza di marche esplicite (Ø). Al contrario, esso non accompagna mai i referenti indefiniti non-specifici né tantomeno funziona da vero e proprio articolo indefinito generalizzato (stadi iv e v). Le testimonianze cinquecentesche possono dunque essere interpretate come la dimostrazione di un processo incipiente di grammaticalizzazione di 'uno' – processo favorito, almeno in parte, dalle strutture morfosintattiche delle lingue modello nelle quali 'uno' si trovava a uno stadio di grammaticalizzazione più avanzato.

Abbreviazioni

1 - 1a persona	INS - strumentale
2 - 2a persona	INT - interrogativo
3 - 3a persona	LOC - locativo
ACC - accusativo	NEG - negazione
ACT - attivo	NOM - nominativo
ADV - avverbio	PASS - passivo
ALL - allativo	PL - plurale
COND - condizionale	PRS - presente
DAT - dativo	PRTC - participio
DEF - definito	PST - passato
DEM - dimostrativo	REFL - riflessivo
DIM - diminutivo	REL - relativo
DU - duale	SG - singolare
F - femminile	SMPRTC - semiparticipio (lit. <i>pusdā-livys</i>), <i>participium contemporale</i>
GEN - genitivo	SN - sintagma nominale
ILL - illativo	SUBJ - congiuntivo
IMP - imperativo	
INF - infinito	

Sigle delle fonti

BB 1563 = BIBlia święta /| Tho ieft /| Księgi Starego y | Nowego zakonu /
własnie z | żydowłkiego / Greckiego y | Ljcińłkiego / nowo na | polski
ięzyk z pilno=|ścią y wiernie | wyłożone. 1563. Edizione fotografica:
<http://polona.pl/item/11638206/6/>.

BrB 1590 = [Jonas Bretkūnas,] BIBLIA | tatai efti|Wišas Schwentas
Raščhtas, Lietuwifchkai pergulditas |per|Janą Bretkuną Lietuwos
pleboną Karaliacziuię |1590. Edizioni critiche: Palionis e Žukauskaitė
(1983, *par.*); Range (2017, *Hrsg.*).

ChB 1660 = [Samuelis Boguslavas Chylinskis,] BIBLIA tatey ira Ra-
sztas Szwėtas seno ir naujo TESTAMENTA. Pirmà kartà dabar per-
guldytas Lietuwiszkañ liežuwiñ. užmariofė. nog Samuelies Baguslawo
CHylinska. Lietuwniko. ó nakłodu Diewo-baymos Karalistes Anglios,
už karalawima Jo Milistos Karalaus, KAROLO ANTROIA Karalaus
Anglios, Szkocios, Irlandios ir Francios. Edizione critica e fotografica:
Kavaliūnaitė (2008, *par.*).

Cos. 1848 = [Giuseppe Costa,] IL | LIBRO DI GIOBBE |VOLGARIZZATO |
PER GIUSEPPE COSTA [...] | ROMA | TIPOGRAFIA DELLA SOCIETÀ EDI-
TRICE ROMANA | [...] |1848.

DP 1599 = [Mikalojus Daukša,] Poštilla | CATHOLICKA. | Tái eft: |
Ižguldimas Ewan|gelii kiekvienos Nedelos ir šwo|tes per wiffūs metūs.
| Per Kúnigą MIKALOIV DAVKSZA | Kanonika Médnikų / iž lėkiško
pergūldita. | Su walá ir dałaidimu wíreufiuių. | W Wilniui /| Drukár-
nioi Akadėmios SOCIETATIS | IESV, A. D. 1599. Edizione critica: Palio-
nis (2000, *par.*). Edizione digitale a cura di Veronika Adamonytė, Milda
Lučinskienė, Jūratė Pajėdienė, Mindaugas Šinkūnas, Eglė Žilinskaitė e
Ona Aleknavičienė: <https://seniejirastai.lki.lt/db.php?source=2>.

Ench. 1561 = [Abel Will,] ENCHIRIDION. | Der Kleine | Catechismus |
Doctor Martin Lu=|thers / Teutſch vnd Preußiſch. | Gedruckt zu Kr-
nigsperg in Preuffen | durch Iohann Daubman. | M. D. LXI. Edizione
digitale e fotografica: <http://www.prusistika.flf.vu.lt>.

Luth. 1545 = [Martin Lutero,] Biblia: das ift: | Die gantze heilige | Schrift:
Deudſch. | Auffſ new zugericht. | D. Mart. Luth. | Begnadet mit
Kur=|fürſtlicher zu Sachſſen Freiheit. | Gedruckt zu Wittem=|berg
/ Durch Hans Lufft. | M. D. XLV. Edizione digitale a cura di Hans

Jürgen Herbst: <http://www.lutherdansk.dk/LutherBiblia1545/biblia2/default.htm#inhalt>.

LXX = Septuaginta. Edizione critica: Martone (2013, a cura di).

MžFK 1559 = [Martynas Mažvydas,] FORMA | Chrikřtima. | Kaip Bařnicřas Iřtati=| mæ Hertzikiftes Prufu/ | ir kitofu temefu lai=|koma ira. | Drukawot Karalaucřui | per Jona Daubmana / | Metu Chriřtaus | M.D.LIX. Edizione fotografica: Subačius (1993, *par.*). Edizione digitale a cura di Diego Ardoino: <https://titus.uni-frankfurt.de/texte/etcs/balt/lit/forma/forma.htm>.

MžG I 1566 = [Martynas Mažvydas,] Gefmes Chrik=|řzoniskas gedomas Bař=|nicřofu per Aduenta ir Ka=|ledas ik Gram=|nicřu. | Iřch řpaufřas Karalau=|cřui nŭg Jona Daubmana | Metu Diewa/| M.D.LXVI. Edizione fotografica: Subačius (1993, *par.*).

MžG II 1570 = [Martynas Mažvydas,] GESMES | Chrikřzonifkas ge=|domas bařnicřofu per | Welikas ir Sekmi=|nias ik Aduenta. | Iřchřpaufřas Kara=|laucřui / nŭg Jona | Daubmana. | Metu Diewa/| M. D. LXX. Edizione fotografica: Subačius (1993, *par.*).

MžK 1547 = [Martynas Mažvydas,] CATE=|CHISMVSA PRA=|řty řzadei, Makřlas řkai=|tima řařhta yr giefmes | del krikřřčianiftes bei del | ber-neliu iaunu nauiey | řugulditas |KARALIAVCZVI VIII. | dena Menefes Saufia, | Metu řřřgimima Diewa. | M. D. XLVII. Edizione fotografica: Subačius (1993, *par.*). Edizione digitale a cura di Diego Ardoino e Jolanta Gelumbeckaitė: <https://titus.fkidg1.uni-frankfurt.de/texte/etcs/balt/lit/mazvkat/mazvk.htm>.

řB 1755 = [Adomas Fridrichas řimelpenigis,] BIBLIA, | Tai eřti: | wiřřas | řřwentas | Rářřtas, | řéno ir Naujo | řeřtamento, | Pagal Wokiřřķą řéřřáttimą | D. Mertino Luterus, | Su | kiekwieno perřkyrimo | trumpu praneřřimmu, | ir | reikalingu pařenklinnimu tu pacřŭ řodřŭ, | kurrie kittoře <...> randomi, | Mokytojŭ Lietuwof | Lietuwifřay řéřřáttitas, | Ir anřřą kartą iřřřpaufřas. | Karaláucřuje, 1755. | Randamas pas Pilippą Kriřtupą Kanerį.

Tal 1764 = [Marcantonio Talleoni,] VOLGARIZZAMENTO | IN TERZA RIMA | DEL SACRO LIBRO | DI GIOB | DEDICATO | A NOSTRO SIGNO-RE | PAPA CLEMENTE XIII | DA | MARCANTONIO TALLEONI | PATRI- ZIO OSIMANO | CAMERIER D'ONORE DI SUA SANTITÀ. | IN OSIMO. CIO IO CC LXIV. | APPRESSO DOMENICANTONIO QUERCETTI. | CON LICENZA DE' SUPERIORI. Edizione fotografica: <https://books>.

google.it/books?id=DMaaCg3os_IC&pg=PA34&hl=it&source=gbs_toc_r&cad=3#v=onepage&q&f=false.

VaišŽT 1600 = [Simonas Vaišnoras,] MARGARITA | THEOLOGICA. | Collecta & confcripta | à | M. ADAMO FRAN=|cisci, Iegerndorfenſe, Mona-|ſterij Heilsbrunnensis Abbate. | Zemcžuga Theologiſchka | Kuri | Sawieie aptur / trumpa parodima / ir | iſchguldima / Wiriaufüüu daliu moxlo Chrik=| ſconifchkoio / pro Bažniczes ir Iſchkallas | Orthodoxiſchkoſes Confefſies | Auguſtanos / alba Sassu. | O nu: | Lituwiſchkai jra perguldita per | Simona Waifchnora warniſchki | Plebona Ragaines | Iſchſpauſta Karaliaucžuiie Pru=|ſu / per Iurgi Ofterbergera / mæ=|tūſa Chriſtaus 1.600. Edizione digitale a cura di Vilma Zubaitienė: <https://seniejirastai.lki.lt/db.php?source=36>.

VE 1579 = [Baltramiejus Vilentas,] ENCHIRIDION | Catechiſmas | maſas / dæl paſpalitu | Plebonu ir Kožnadiju / | Wokiſchku lieſuwiu para=|ſchits per Daktara Mar=|tina Luthera. | O iſch Wokiſchka lieſuwia ant | Lietuwiſchka pilnai ir wiernai pergul=|dita / per Baltramieju Willentha | Plebona Karalaucžuiie ant | Schteindama. | Iſchſpauſtas Karalau=|cžui per Iurgi Ofterber=|gera / Metu Diewa | M. D. LXXIX. Edizione digitale a cura di Ona Aleknavičienė: <https://seniejirastai.lki.lt/db.php?source=87>. Edizione critica e fotografica: Ford (1969, ed.).

VEE 1579 = [Baltramiejus Vilentas,] Euangelias bei | Epiſtolas / Nedeliu ir ſchwen=|tuju dienofu ſkaitomofias / Bažnicžo=|ſu Chrikſczonifchkoſu / pilnai ir wiernai per=|guldita ant Lietuwiſchka Szodžia / per |Baltramieju Willenta / Plebo=|na Karalaucžui ant |Schteindama. | Priegtam / ant gala priedeta jra Hiſto=|ria apie muka ir ſmerti Wiefchpaties muſu |Ieſaus Chriſtaus pagal kieturiu |Euangeliftu. |Iſchſpauſtas Karalaucžui per Iurgi | Ofterbergera / Metu |M. D. LXXIX. Edizione digitale a cura di Ona Aleknavičienė: <https://seniejirastai.lki.lt/db.php?source=88>.

Vulg = Bibliorum Sacrorum Nova Vulgata Editio: https://www.vatican.va/archive/bible/nova_vulgata/documents/nova-vulgata_index_lt.html.

WP 1573 = [Anonimo,] ISCHGVLDIMAS | EVANGELIV PER WISVS MET=|TVS, SVRINKTAS DALIMIS ISCH | DAVGIA PASTILLY, TAI EST | ISCH PASTILLAS NICVLAI HE=|MINGY, ANTONY CORVINI, IO=| ANNIS SPANGENBERGI, MAR=|TINI LVTHERI, PHILIPPI | MELANTHONIS, IOANNIS | BRENTY, ARSATY, SCHO=|PER, LEONARDI KVLMA=|NI IODOCY WILICHI | IR ISCH KIT=|TV ·ϕ·. Edizione digitale a cura di Jolanta Gelumbeckaitė: <https://seniejirastai.lki.lt/db.php?source=37>; <https://titus.fkidg1.uni-frankfurt.de/texte/etcs/balt/lit/wp/wp.htm>.

Bibliografia

- ALEKSYNAS, K. (1985, *par.*), *Stebuklingas žodis. Lietuvių liaudies pasakos*, Šviesa, Kaunas.
- ALEW = HOCK, W. *et al.* (2005, *Hrsg.*), *Altltauisches etymologisches Wörterbuch*. 2. Band: *N-Ž*, Baar, Hamburg.
- BAZZANELLA, C. (2011), *Approssimare con i numeri nelle lingue*, in BAZZANELLA, C., PUGLIESE, R. e STRUDSON, E. (2011, a cura di), *Numeri per parlare. Da 'quattro chiacchiere' a 'grazie mille'*, Laterza, Bari, pp. 21-58.
- CARDINALETTI, A. e GIUSTI, G. (2018), *Indefinite Determiners: Variation and Optionality in Italo-Romance*, in D'ALESSANDRO, R. e PESCARINI, D. (2018, eds.), *Advances in Italian Dialectology. Sketches of Italo-Romance Grammars*, Brill, Leiden, pp. 135-161.
- CERRI, A. (2016), *Interferenza basso-tedesca nei testi lettoni antichi: l'espressione della (in)definitezza*, in BENEDETTI, M., BRUNO, C., DARDANO, P. e TRONCI, L. (2016, a cura di), *Grammatiche e grammatici. Teorie, testi e contesti. Atti del XXXIX Convegno della Società Italiana di Glottologia*, Il Calamo, Roma, pp. 219-225.
- CREISSELS, D. (2019), *Inverse locational predication in typological perspective*, in «Italian Journal of Linguistics», 31, 2, pp. 38-106.
- CREISSELS, D., BERTINETTO, P.M. e CIUCCI, L. (in stampa), *Non-verbal predication in the world's languages: An overview*, in BERTINETTO, P.M., CIUCCI, L. e CREISSELS, D. (in stampa, eds.), *Non-verbal predication in the world's languages: A typological survey* (Comparative Handbooks of Linguistics 9).
- DE LA VILLA, J. (2010), *Numerals*, in BALDI, P. e CUZZOLIN, P. (2010, eds.), *New Perspectives on Historical Latin Syntax*. Vol. 3: *Constituent Syntax: Quantification, Numerals, Possession, Anaphora*, De Gruyter Mouton, Berlin / New York, pp. 175-238.
- DINI, P.U. (1997), *Le lingue baltiche*, La Nuova Italia, Firenze.
- DINI, P.U. (2014), *„ins undeudſche gebracht“. Sprachgebrauch und Übersetzungsverfahren im altpreussischen Kleinen Katechismus*, De Gruyter, Berlin / Boston.

- DRYER, M.S. (2013), *Indefinite Articles*, in DRYER, M.S. e HASPELMATH, M. (2013, eds.), *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig [accessibile online: wals.info/chapter/38].
- FORD, G.B. JR. (1969, ed.), *The Old Lithuanian catechism of Baltramiejus Vilentas (1579): a phonological, morphological and syntactical investigation*, Mouton, The Hague / Paris.
- GEIST, L. (2013), *Bulgarian Edin: The Rise of an Indefinite Article*, in JUNGHANNS, U., FEHRMANN, D., LENERTOVÁ, D. e PITSCH, H. (2013, eds.), *Formal Description of Slavic Languages: The Ninth Conference*, Peter Lang, Frankfurt am Main, pp. 125-148.
- GIVÓN, T. (1981), *On the Development of the Numeral 'One' as an Indefinite Marker*, in «Folia Linguistica Historica», 2, pp. 35-53.
- HASPELMATH, M. (1997), *Indefinite Pronouns*, Oxford University Press, Oxford.
- HEINE, B. (1997), *Cognitive Foundations of Grammar*, Oxford University Press, New York / Oxford.
- HOLVOET, A. e TAMULIONIENĖ, A. (2006), *Apibrėžtumo kategorija*, in HOLVOET, A. e MIKULSKAS, R. (2006, par.), *Daiktavardinio junginio tyrimai*, Lietuvių kalbos institutas, Vilnius, pp. 11-32.
- KAUNAS, D., KUOLYS, D., KOŽENIAUSKIENĖ, R., LANDSBERGIS, V., TRILUPAITIENĖ, J. e TYLA, A. (1998, par.), *Martynas Mažvydas and Old Lithuania*, Pradai, Vilnius.
- KAVALIŪNAITĖ, G. (2008, par.), *Samuelio Boguslavo Chylinskio Biblija. Senasis Testamentas*. Vol. 1, Lietuvių kalbos institutas, Vilnius.
- KRIFKA, M. (2007), *Approximate Interpretation of Number Words: A Case for Strategic Communication*, in BOUMA, G., KRÄMER, I. e ZWARTS, J. (2007, eds.), *Cognitive Foundations of Interpretation*, Royal Netherlands Academy of Arts and Sciences, Amsterdam, pp. 111-126.
- KUDZINOWSKI, C. (1977, ed.), *Indeks - słownik do "Daukšos Postilė"*. Vol. 2: O-Ž, UAM, Poznań.
- KUTEVA, T., HEINE, B., HONG, B., LONG, H., NARROG, H. e RHEE, S. (2019), *World Lexicon of Grammaticalization. 2nd edition*, Cambridge University Press, Cambridge.

- LAMBRECHT, K. (2000), *When subjects behave like objects: an analysis of the merging of S and O in sentence-focus constructions across languages*, in «Studies in Language», 24, 3, pp. 611-682.
- LKŽe = *Lietuvių kalbos žodynas*. 20 voll. (1941-2002). Versione elettronica a cura di NAKTINIENĖ, G., PAULASKAS, J., PETROKIENĖ, R., VITKAUSKAS, V. e ZABARSKAITĖ, J., Lietuvių kalbos institutas, Vilnius [accessibile online: www.lkz.lt].
- LÚJÁN MARTINEZ, E.R. (1995), *Los usos no numerales de las palabras para «uno» y la raíz indoeuropea *sem*, in «Cuadernos de Filología Clásica», 5, pp. 215-230.
- MARTONE, C. (2013, a cura di), *La Bibbia dei Settanta*. Vol. 3: *Libri poetici*, Morcelliana, Brescia.
- MCGREGOR, W.B. (2014), *Numerals and number words in Shua*, in «Journal of African Languages and Linguistics», 35, 1, pp. 45-90.
- MCNALLY, L. (2019), *Existential sentences*, in PORTNER, P., MAIENBORN, C. e VON HEUSINGER, K. (2019, eds.), *Semantics*. Vol. 5: *Sentence and information structure*, Mouton De Gruyter, Berlin / Boston, pp. 281-305.
- MEISTERFELD, R. (2000), *Die unbestimmte Bestimmung: Zur Entstehung des unbestimmten Artikels in den romanischen Sprachen*, in STAIB, B. (2000, Hrsg.), *Linguistica romanica et indiana: Festschrift für Wolf Dietrich zum 60. Geburtstag*, Narr, Tübingen, pp. 303-332.
- MICHELINI, G. (2000, par.), *Martyno Mažvydo raštai ir jų šaltiniai*, Mokslo ir enciklopedijų leidybos institutas, Vilnius.
- MICHELINI, G. (2002), *Gli inni sacri dei lituani di Prussia. Da Mažvydas a Kleinas*, Monte Università Parma editore, Parma.
- MOLINARI, L. (2022), *The Syntax of Polish jeden 'One' as an Indefinite Determiner*, in «Annali di Ca' Foscari. Serie occidentale», 56, pp. 63-84.
- NAPOLI, M. (2013), *When the indefinite article implies uniqueness: A case study from Old Italian*, in «Folia linguistica», 47, 1, pp. 183-235.
- PALIONIS, J. e ŽUKAUSKAITĖ, J. (1983, par.), *Jonas Bretkūnas. Rinktiniai raštai*, Mokslo, Vilnius.
- PALIONIS, J. (2000, par.), *Mikalojaus Daukšos 1599 metų Postilė ir jos šaltiniai*, Baltos lankos, Vilnius.

- PARENTI, A. (1995), *Note sul problema dell'articolo prussiano*, in «Ponto-Baltica», 6, pp. 71-87.
- RANGE, J.D. (2017, Hrsg.), *Das Neue Testament. Evangelien und Apostelgeschichte. Labiau 1580*, Verlag Ferdinand Schöningh, Paderborn.
- SCHROEDER, C. (2006), *Articles and article systems in some areas of Europe*, in BERNINI, G. e SCHWARTZ, M.L. (2006, eds.), *Pragmatic organization of discourse in the languages of Europe*, Mouton de Gruyter, Berlin / New York, pp. 545-611.
- SMOCZYŃSKI, W. (2018), *Lithuanian Etymological Dictionary*, Peter Lang, Berlin.
- SPRAUNINENĖ, B. (2008), *Apibrėžtumo kategorija ir jos tyrimai lietuvių kalbotyroje*, in «Lietuvių kalba», 2 [accessibile online: <https://www.zurnalai.vu.lt/lietuviu-kalba/article/view/22893/22170>].
- SUBAČIUS, G. (1993, par.), *Martynas Mažvydas. Katekizmas ir kiti raštai / Catechismus und andere Schriften*, Baltos lankos, Vilnius.
- URBAS, D. (1998 [1996¹]), *Martyno Mažvydo raštų žodynas*, Mokslo ir enciklopedijų leidybos institutas, Vilnius.
- WHERRITY, M. (2016), *Ex uno plures: A case for monosemy*, in «Nordic Journal of English Studies», 15, 2, pp. 202-214.
- WIESE, H. (2003), *Numbers, Language, and the Human Mind*, Cambridge University Press, Cambridge, MA.
- WIESE, H. (2007), *The Co-Evolution of Number Concepts and Counting Words*, in «Lingua», 117, 5, pp. 758-772.
- ZINKEVIČIUS, Z. (1996), *The History of the Lithuanian Language*, Mokslo ir enciklopedijų leidykla, Vilnius.

ADRIANO CERRI

Dipartimento di Filologia, Letteratura e Linguistica

Università di Pisa

Via Santa Maria 36

56126 Pisa (Italia)

adriano.cerri@unipi.it

Il trattamento di /w/ in posizione iniziale in italiano è influenzato dall'ortografia? Indizi dalla selezione dell'articolo maschile

PIERO COSSU

ABSTRACT

Italian native speakers show two patterns concerning the selection of the singular masculine definite article allomorph before a glide /w/ or a sibilant + glide cluster /sw/. As far as Italian native lexemes are concerned, speakers select *l'* and *il* before /w/ and /sw/, respectively. Differently, *il* and *lo* are selected before loanwords showing the same contexts in first position. We propose that the observed divergence is due to two syllabifications of the phonemic strings including a glide /w/. In turn, the different syllabifications depend on the graphemic representation of the glide: <u> in the native lexicon, mostly <w> in loanwords. It is thus proposed that the vocalic versus consonantal status of the grapheme representing the glide influences the phonological representation of the glide itself. If /w/ is written as <u>, it occupies a nuclear position, if /w/ is written as <w>, it stands in onset position. We conducted an analysis on the vast Google Books corpus to corroborate this view. Moreover, a phonological analysis of Italian diphthongs is discussed.

KEYWORDS: diphthong, article allomorph, glide, orthography.

1. Introduzione

La distribuzione degli allomorfi dell'articolo maschile singolare in italiano è, com'è noto, determinata in gran parte dalla struttura sillabica nella posizione iniziale della parola seguente (Kaye, Lowenstamm e Vergnaud, 1990; Marotta, 1993; Petrosino, 2018). I tre allomorfi *il*, *lo* e *l'* sono infatti in distribuzione complementare secondo il seguente schema:

- *il* ricorre davanti ad attacchi sillabici, siano essi semplici o ramificati: *il cane*, *il treno*;

- *lo* si seleziona davanti a gruppi coda-attacco, rappresentati da *sC* o dalle cosiddette geminate intrinseche /p:/, /f:/, /tʃ:/, /dʒ:/ e, anche se molto più rara in prima posizione, /k:/: *lo stato*, *lo sgarro*, *lo gnomo*, *lo sciocco*, *lo zio*, *lo zero*, *lo gliommero*; anche *j* seleziona di norma *lo*: *lo yogurt*, *lo iato*¹;
- *l'* è usato invece di fronte a nuclei, semplici o ramificati: *l'atomo*, *l'olmo*, *l'uomo*, *l'uovo*.

Riguardo a quest'ultimo contesto, ossia un dittongo ascendente in posizione iniziale il cui primo elemento è un'approssimante labiovelare /w/, tuttavia, è stata più volte osservata una duplice possibilità. Infatti, se nel lessico nativo, scarso per questo contesto fonologico, si seleziona l'allomorfo consonantico *l'* (*l'uomo*, *l'uovo*), in certi prestiti si tende a selezionare l'allomorfo *il* (Marotta, 1993; Bertinetto e Loporcaro, 2005; Canalis, 2018; Baroni, 2020), come ad esempio in *il whisky*, *il western*.

Inoltre, si osserva un'ambiguità analoga nel caso della sequenza costituita da sibilante più approssimante labiovelare /sw/. Nel lessico nativo l'allomorfo selezionato è *il* (*il suocero*, *il suono*), mentre nei prestiti la tendenza è verso *lo* (*lo Swatch*, *lo swing*), cfr. Janni (1992); focus di questo studio è la selezione dell'articolo davanti alle parole che iniziano con il legamento /w/ o con il nesso /sw/.

In linea con l'idea avanzata, con cautela, da Camilli e Fiorelli (1965: 194) e difesa più solidamente da Janni (1992) e Baroni (2020), riteniamo che le ambiguità illustrate sopra possano essere governate dalla diversa rappresentazione ortografica del legamento (<u> o <w>). La nostra ipotesi è che vi sia una diversa sillabificazione del dittongo *wV* dipendente dalla forma ortografica del legamento e che la corrispondente selezione dell'allomorfo dell'articolo sia un epifenomeno delle due diverse sillabificazioni. In sintesi, il dittongo *wV* occuperebbe fonologicamente la posizione di nucleo quando /w/ è scritto <u>, mentre sarebbe assimilato ad una sequenza CV, quindi attacco e nucleo, se /w/ è scritto <w>.

¹ Come vedremo, di fronte a /j/ tuttavia la situazione è soggetta a maggior variazione. *Gliommero* è un prestito dal napoletano [ʎjom:ero] "gomitolo" usato in italiano per indicare un tipo di componimento poetico.

Si vedrà dapprima, nel paragrafo 2, una rassegna critica degli studi fonologici che si sono occupati dei dittonghi in italiano, sia in prospettiva diacronica che sincronica. Nel paragrafo 3 valuteremo la distribuzione degli allomorfi dell'articolo maschile singolare in italiano davanti a parole inizianti con /w/ e /sw/ in riferimento sia a lessemi nativi che a prestiti. Nel paragrafo 4 prenderemo in esame gli adattamenti fonologici dei prestiti, in particolare di quelli inglesi in italiano. Nei paragrafi 5 e 6 presenteremo dati originali, concernenti la selezione dell'articolo davanti a /j/ e /w/, ottenuti tramite lo strumento *Ngram Viewer* di *Google*, che permette di consultare con relativa semplicità l'enorme corpus offerto da *Google Libri*. La discussione dei dati è affrontata nel paragrafo 7.

2. I dittonghi in italiano

Con 'dittongo' si intende una sequenza tautosillabica di due vocali adiacenti. Esso si oppone a 'iato', termine che indica una sequenza vocalica eterosillabica (Marotta, 2010). Se da un punto di vista sopra-segmentale la differenza riguarda principalmente il numero di sillabe coinvolte nella sequenza vocalica, al livello segmentale la differenza risiede nel fatto che i dittonghi, a differenza degli iati, presentano un'approssimante in luogo di uno dei due membri vocalici costitutivi; gli iati sono infatti caratterizzati da due vocali piene contigue. Per l'italiano, non c'è unanime consenso riguardo alla posizione nucleare o extranucleare dell'approssimante presente nei dittonghi. Prima di procedere, occorre innanzitutto distinguere tra dittongo ascendente (GV)² e dittongo discendente (VG). Se G precede la vocale si ha un dittongo ascendente e il legamento è definito come semiconsonante /w j/, mentre se la vocale è seguita da un'approssimante il dittongo è discendente e G è descritto come semivocale /ɨ ʉ/. A parte questa prima differenziazione di tipo posizionale, si riscontrano importanti differenze fonetiche tra GV e VG, poiché aspetti temporali e timbrici distinguono i due tipi di dittongo. In Salza (1988: 104) si riportano durate sensibilmente

² 'G' indica l'approssimante o legamento, dall'inglese *glide*, 'V' sta per la vocale piena.

inferiori per le semiconsonanti rispetto alle semivocali, sia in sillaba tonica che atona. Lo studio di Salza, Marotta e Ricca (1987) ha messo in luce anche differenze dal punto di vista formantico: se le semivocali sono simili alle vocali atone, per i valori di F2, le semiconsonanti tendono invece ad avvicinarsi al centro dello spazio vocalico. Inoltre, nelle semiconsonanti, generalmente più brevi, è pressoché assente una fase di tenuta, poiché si tratta di segmenti altamente dinamici.

In italiano, strutture di tipo GV traggono origine da diversi contesti. I dittonghi ascendenti con approssimante labiovelare [w] (d'ora in avanti *wV*) possono derivare dal mantenimento delle labiovelari latine (Marotta, 2010), come in *acqua, quando, lingua, sangue*, oppure da germanismi, i quali hanno sviluppato un'occlusiva velare sonora in attacco, come *guerra* < **werra* (Cortelazzo e Zolli, 1980). I dittonghi ascendenti con approssimante palatale [j] (d'ora in poi *jV*) possono derivare dall'evoluzione in legamento di una vocale alta o media, originariamente il primo elemento di uno iato ([*'rab:ja*] < *RABIA*), o della laterale [l] nei gruppi *muta cum liquida* di etimi latini, come in [*'pjanta*] < *PLANTAM* (Marotta, 2010). Inoltre, sia *wV* che *jV* possono derivare dalla dittongazione romanza, che in italiano (o, meglio, toscano) ha interessato le vocali medie brevi latine in sillaba aperta (Serrianni, 1998: 47; Loporcaro, 2011: 120): *BŎNUM* > [*'bwɔ:no*] *buono*, *LŎCUM* > [*'lwɔ:go*] *luogo*, *HŎMO* > [*'wɔ:mo*] *uomo*, *PĚDEM* > [*'pjɛ:de*] *piede*, *PĚTRAM* > [*'pjɛ:tra*] *pietra*, *HĚRI* > [*'jɛ:ri*] *ieri*.

Come si evince dal quadro sopra riportato, *wV* ha maggiori restrizioni distribuzionali rispetto a *jV*. A partire da Marotta (1987) è noto in letteratura che [j] può precedere qualsiasi vocale (tranne [i]), ad esempio [*'jɛ:na*] *iena*, [*'jɔ:djo*] *iodio*, [*'ja:to*] *iato*, [*'ju:ta*] *iuta*, e seguire qualsiasi consonante ([*'bjaŋko*] *bianco*, [*'kjɛ:za*] *chiesa*, [*'sjɛ:ro*] *siero*, [*'mjɛ:le*] *miele*, [*'fja:la*] *fiala*, ecc.). Diversamente, [w] può essere preceduto da qualsiasi consonante, come in [*'bwɔ:no*] *buono*, [*'lwɔ:go*] *luogo*, [*'twɔ:no*] *tuono*, [*'swɔ:ra*] *suora*, [*'rwɔ:ta*] *ruota*, ecc., mentre può essere seguito da vocali diverse da [ɔ] (tranne l'omorganica [u]) solo quando è preceduto da occlusiva velare [k g]: [*'kwa:le*] *quale*, [*'kwel:o*] *quello*, [*'kwɛstua*] *questua*, [*'kwindi*] *quindi*, [*'gwanto*] *quanto*, [*'gwɛrtfo*] *guercio*, [*'gwi:za*] *guisa*.

In italiano, attualmente, i prestiti recenti che recano in posizione iniziale *wV*, dove *V* ≠ [ɔ], non presentano la prostesi di [g] come avveniva in latino tardo³ e i cui esiti si possono riscontrare anche in spagnolo e francese: *wardan* > sp. [gwarðar] (Jacobs e Van Gerwen, 2006: 78ss.). Nel trattamento di prestiti dall'inglese, la prostesi di [g] davanti a #/w/ sembra essere un fenomeno ancora produttivo in spagnolo, come in [har(d)gwer] *hardware*, [sof(t)gwer] *software* (De la Cruz Cabanillas *et al.*, 2007: 59), *windshield* > [gwintʃil], *to weld* [gweldear] (per lo spagnolo chicano, Smead, 1998: 116, 118), fenomeno condiviso sempre in prestiti dall'inglese anche dal chamorro, lingua maleo-polinesiana parlata nelle Isole Marianne e Guam (Witucki, 1976: 13).

Per tornare all'italiano, anche formazioni interne come UIL [wil], acronimo per Unione Italiana del Lavoro, risultano perfettamente accettabili (Canalis, 2018: 19). Inoltre, parallelamente a prestiti come *Twitter*, (*Kinder*) *Bueno* (Canalis, 2018), un gran numero di italofofoni L1 può realizzare dittonghi di tipo *wV* non preceduti da occlusiva velare e seguiti da vocali diverse da [ɔ] in parole come [pun'twa:le] *puntuale*, [at'twa:le] *attuale*, [kon'ti:nwi] *continui*, [swa'dente] *suadente*, [as:we'fa:re] *assuefare*, [per'swa:zo] *persuasore*⁴.

2.1. La sillabificazione dei dittonghi ascendenti

Possiamo ora tornare alla questione iniziale, ossia quale posizione occupano i legamenti all'interno della sillaba: nucleare o extranucleare? Per provare a rispondere alla domanda si metteranno a confronto

³ Il fenomeno è ancora produttivo nel lessico nativo di molti dialetti spagnoli: *bueso* "osso" /weso/ → [ˈgweso] (BRADLEY, 2014: 67), *hueco* "buco" → [ˈgweko] (HUALDE, 1989: 825), *huevo* "uovo" /ue.bo/ → [ˈgwe.βo] (MORALES-FRONT, 2018: 203). Si noti che tutti gli esempi riguardano un'approssimante labiovelare seguita dalla vocale media anteriore. Inoltre, l'*Observatorio de Palabras* della *Real Academia Española* ritiene valide anche nello scritto le forme *guásap* (<https://www.rae.es/observatorio-de-palabras/guasap>), indicante il messaggio di testo inviato tramite il noto servizio di messaggistica istantanea *WhatsApp*, *guasapear*, l'atto di inviare messaggi in questa modalità (<https://www.rae.es/observatorio-de-palabras/guasapear>), e *güisqui*, dall'inglese *whiskey* (<https://dle.rae.es/g%C3%BCisqui#8dVxzER>).

⁴ MAROTTA (1987: 874) ritiene che in questi casi il dittongo sia possibile, ma che lo iato sia favorito nei casi in cui la sequenza sia tonica, mentre VAN DER VEER (2006: 78) e CANALIS (2018: 19) non pongono questa restrizione.

principalmente le proposte di Marotta (1987; 1988) e di Canalis (2018), che sono le più esaustive offerte in letteratura.

Marotta (1987; 1988) distingue innanzitutto *wV* da *jV*, da un punto di vista fonologico. Nell'ottica dell'autrice, l'approssimante palatale [j] appartiene all'attacco, come segmento autonomo. Diversamente, [w] è, nel caso di *wV* preceduto da consonante velare, un'articolazione secondaria delle occlusive [k g], interpretabili quindi come occlusive labiovelari [k^w g^w], mentre è parte del nucleo negli altri casi (come in *uomo*, *luogo*). Si vedano le rappresentazioni nella Figura 1.

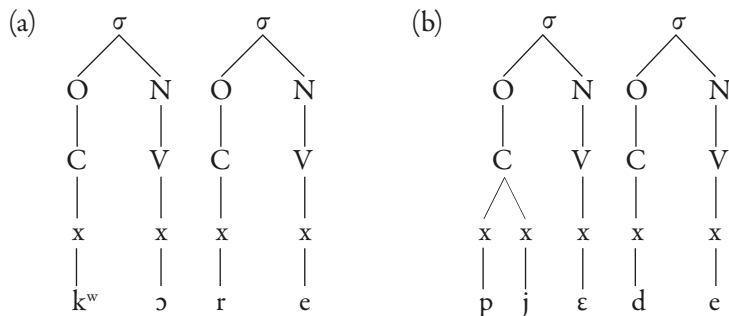


Figura 1. *Rappresentazione sillabica di cuore e di piede in cui Cw corrisponde a un segmento complesso (1a) e Cj a un attacco complesso (1b).*

Le due strutture sillabiche nella Figura 1 non creano particolari problemi teorici e permettono di tenere conto dell'occorrenza di nuclei bimoraici, ossia costituiti da una vocale lunga (in sillaba tonica aperta) o da una vocale e una consonante in coda (primo elemento di gruppi consonantici eterosillabici o geminate). Tuttavia, rappresentazioni come quella nella Figura 1b comportano l'ammissione di costituenti trisegmentali come in *prò.prjo*, in sillaba atona, e *ap.pro.prjår.si*, in sillaba tonica. La rappresentazione in 1a, con occlusiva labiovelare [k^w], proposta in Marotta (1988), che implica di non considerare la stringa *wV* in, ad esempio, *quando* come un dittongo, è accettata anche da Canalis (2018). Nei casi in cui la testa dell'attacco è invece una consonante non velare, sia Marotta (1987; 1988) che Canalis (2018) escludono la possibilità che le approssimanti [j] e [w] possano essere articolazioni secondarie della consonante in attacco.

Nella Figura 2 è illustrata invece la rappresentazione in cui [w] è un membro del nucleo (cfr. Marotta, 1987; 1988).

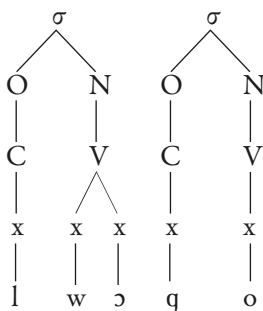


Figura 2. *Rappresentazione sillabica di luogo in cui il dittongo è associato al nucleo.*

Tra i sostenitori della struttura sillabica in Figura 2 vi è anche Van der Veer (2006), che però la ritiene valida solo per le sillabe toniche. Interpretare il dittongo come parte del nucleo permette di tenere conto della struttura sillabica, ad esempio, in ['trwɔ:golo] *truogolo*, ['flwɔ:ro] *fluoro*, [des'trje:ro] *destriero*, [co'prja:mo] *copriamo*. Marotta (1987: 875) riconosce la possibilità che in strutture costituite da *muta cum liquida* + legamento + vocale possa occorrere un dittongo, soprattutto nel parlato trascurato o ad alta velocità, nonostante esista una generale tendenza a favorire gli iati, a causa della presenza di una consonante rotica che precede il legamento.

I problemi con strutture siffatte emergono considerando forme come *tworlo* e *appropriarsi*. In casi del genere a essere trisegmentale sarebbe infatti la rima, interna alla radice nel primo esempio, originata dall'incontro di morfemi nel secondo. Nella prospettiva in cui il legamento non è parte dell'attacco (vedi *supra*), casi come quelli appena illustrati corroborerebbero la proposta di Canalis (2018: 13), secondo cui il dittongo ascendente è da interpretare come una vocale complessa. In altre parole, si tratterebbe di una vocale con due gesti articolatori, una sorta di produzione dinamica, ma monoposizionale. Sarebbe il corrispettivo nucleare di una consonante affricata: come per le affricate, i due gesti articolatori sono disposti in ordine crescente di sonorità e in ordine decrescente di forza.

La rappresentazione di una vocale complessa è data in Figura 3 (Canalis, 2018: 14).

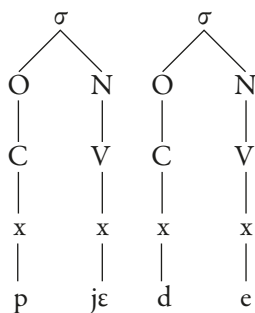


Figura 3. *Rappresentazione sillabica di un dittongo ascendente come segmento complesso.*

La struttura in Figura 3 è stata proposta da Canalis (2018) per le sillabe sia atone che toniche, mentre Van der Veer (2006) la ritiene valida solo per le sillabe atone. Canalis (2018) giunge a questa conclusione considerando, tra gli altri fattori, l'esistenza di trittonghi, strutture trisegmentali ascendenti-discendenti come in [mjeɪ̯] *miei*, [twoɪ̯] *tuoi*, [ma'jeɪ̯tica] *maieutica* e la struttura GGV in [ma'ɾjwɔlo] *mariuolo*. L'autore può interpretare queste strutture come dittonghi nella misura in cui la vocale piena e la semiconsonante a essa adiacente costituiscono una vocale complessa.

Postulare l'esistenza di vocali complesse monoposizionali può essere una delle soluzioni meno problematiche nel trattare la rima tonica negli esempi visti in precedenza: [ˈtwɔrlo] *tuorlo*, [ap:roˈprjarsi] *appropriarsi*. Optando per la biposizionalità incorriamo infatti in costituenti trisegmentali, a discapito del binarismo.

Se si segue la strada della monoposizionalità, nei lessemi come [ma'ɾjwɔlo] *mariuolo* e [a'jwɔla] *aiuola* da una parte, con legame anteriore e posteriore in successione, e dall'altra [se'gwjamo] *seguiamo*, [re'li:kwja]⁵ *reliquia*, [ef:etˈtwjamo] *effettuiamo*, [at:eˈnwjamo] *attenuiamo* con le approssimanti in ordine invertito, le stringhe GGV devono

⁵ In questi casi, MAROTTA (1988) e CANALIS (2018) parlano di consonanti labiovelari (/k^w/, /g^w/).

essere considerate dittonghi biposizionali. Il dizionario Treccani⁶ indica *mariuolo* come quadrisillabico ([mari'wɔ:lo]), pronuncia che in effetti appare meno marcata di quella con trittongo, forse per via dell'attacco con consonante rotica /r/ (cfr. Marotta, 1987: 875). Per *aiuola* invece lo stesso dizionario⁷ propone il trittongo /jwɔ/ (→ [ɥwɔ] o [jɥɔ]), mentre Marotta (1987: 881) oltre al trittongo riconosce due dittonghi contigui ([aj.wɔ]) come possibile alternativa. Si tratta di strutture spesso ignorate dagli studi fonetici, ma non mancano tuttavia analisi mirate a queste specifiche produzioni. Per l'italiano, una di queste è Calamai e Bertinetto (2006). Gli autori, analizzando dati acustici e elettropalatografici, osservano che le due approssimanti contigue in un supposto trittongo sono spesso soggette a coalescenza, emergendo come approssimante labiopalatale [ɥ], come in [akkɥe'ta:re] *acquietare*.

Da un punto di vista fonologico, le produzioni discusse in Calamai e Bertinetto (2006) sono compatibili con l'idea dei dittonghi come vocali complesse (Canalis, 2018)⁸. Se l'approssimante più prossima alla vocale piena è parte integrante della vocale, quella osservata spettrograficamente sarebbe la resa fonetica di un dittongo in cui l'elemento sillabico è una vocale complessa. In quest'ottica, il dittongo ha come primo elemento il legamento labiovelare, che qui emerge nella forma del suo allofono contestuale [ɥ] a causa della coarticolazione con la palatalità del *glide* in [ja] (o [j̞a]), come in *contin*[ɥja]mo (Calamai e Bertinetto, 2006: 53).

Possiamo ora provare ad applicare l'ipotesi della monoposizionalità ai contesti di specifico interesse in questo studio, ossia i dittonghi ascendenti wV. Attuare ciò implica una loro rianalisi in termini di vocali complesse, nuclei monoposizionali (Canalis, 2018). Lessemi come *uomo* saranno quindi trisegmentali /w̃mo/ e *suono* quadrisegmentali /sw̃no/.

⁶ Cf. <https://www.treccani.it/vocabolario/mariolo/>.

⁷ Cf. <https://www.treccani.it/vocabolario/aiuola/>.

⁸ Le stesse produzioni fonetiche, caratterizzate da coalescenza delle due approssimanti /wj/ → [ɥ], sono compatibili anche con altre interpretazioni dei dittonghi, ossia come strutture biposizionali. Le produzioni relative ai trittonghi non devono essere intese come evidenza della monoposizionalità dei dittonghi. Esse suggeriscono tuttavia che l'ottica monoposizionale non è incompatibile con la scenario mostrato dai trittonghi.

Applicata ai *wV* tonici, l'ipotesi della monoposizionalità ha il vantaggio di tener conto dell'allungamento vocalico in sillaba tonica aperta non finale (Marotta, 1985) che caratterizza l'italiano. La vocale tonica in *uomo*, *uovo*, *suono* è infatti lunga ([*wɔ:mo*], [*wɔ:vo*], [*swɔ:no*]) e le vocali lunghe occupano due posizioni nella struttura scheletrica. Se i dittonghi devono essere considerati come nuclei complessi, allora il nucleo in queste parole è proiettato su tre posizioni. Il problema è risolto ammettendo che i dittonghi di tipo ascendente sono in realtà vocali complesse.

Come vedremo nella prossima sezione, esistono delle situazioni in cui l'interpretazione nucleare dei dittonghi *wV*, sia in ottica biposizionale (Marotta, 1987; 1988) che monoposizionale (Canalis, 2018) è ostacolata dallo scenario della selezione dell'articolo davanti a (*s*)*wV*. Il quadro che presenteremo di seguito impone il riconoscimento di una netta distinzione fonologica tra produzioni fonetiche identiche. Parliamo dei dittonghi con approssimante labiovelare (*wV*) provenienti da prestiti. Come anticipato nel paragrafo 1, i lessemi nativi che iniziano con *wV* selezionano l'allomorfo *l'*, mentre quelli che iniziano con *swV* prevedono *il*. Diversamente, nei prestiti, l'allomorfo selezionato davanti a *wV* è *il* e davanti a *swV* è *lo*.

3. La selezione dell'articolo davanti a (*s*)*wV*

Osservare lo schema di selezione dell'articolo maschile al fine di determinare l'etero- o tautosillabicità di alcuni nessi è una diagnostica usata molto di frequente in letteratura (Davis, 1990; Kaye, 1992; Nespor, 1993; Marotta, 1993; Canalis, 2018), anche se non è esente da critiche. Ad esempio, Canalis (2018) ritiene che l'osservazione dello schema distribuzionale degli allomorfi dell'articolo determinativo maschile non possa avere dal punto di vista sincronico la stessa efficacia che ha avuto in prospettiva diacronica, ossia per l'italiano antico. L'autore fa appello a specifiche questioni riguardanti proprio il legamento [w] in posizione iniziale, che selezionerebbe *l'* o *il* in maniera apparentemente casuale. Come anticipato nel paragrafo 1, riteniamo

invece che sia possibile superare la presunta ambiguità nella selezione dell'articolo in italiano davanti ai dittonghi.

I lessemi del lessico nativo italiano che iniziano con [w] sono due, *uomo* e *uovo*, e selezionano l'allomorfo *l'*⁹. Al contrario, i prestiti selezionano *il*: *il western*, *il web*, *il weekend* (Bertinetto e Loporcaro, 2005; Canalis, 2018; Baroni, 2020). Parallelamente, le sequenze *swV* del lessico nativo selezionano *il* (*il suono*, *il suocero*), ma i prestiti con la stessa stringa iniziale (/sw/) selezionano *lo*: *lo Swatch*, *lo swing*, *lo swahili*¹⁰ (Janni, 1992; Serianni, 1998).

Per quanto riguarda *wV*, si potrebbe fare appello al fatto che l'italiano non ammette dittonghi diversi da /wɔ/ se la consonante precedente non è un'occlusiva velare. È possibile infatti che l'ambiguità dello schema della selezione dell'articolo sia motivata da una differente analisi fonologica dei dittonghi in questione sulla base del timbro di V. La selezione di *il* e di *lo*, davanti ai casi in cui la vocale in *wV* e *swV* non sia /ɔ/, potrebbe quindi essere dovuta a un'interpretazione della struttura sillabica del dittongo in cui il legamento è in posizione di attacco. In quest'ottica, i nessi *swV* con V ≠ /ɔ/ sarebbero quindi reinterpretati come *sC*. Come tuttavia fanno notare Bertinetto e Loporcaro (2005: 142), recenti prestiti con /wɔ/ selezionano comunque *il* (*il* [wɔ] *lkie-talkie*) e, come illustrato da *Swatch*, quelli con /swɔ/ selezionano *lo*, se prestiti. Si potrebbe dire che la deviazione dalla norma sia una caratteristica che riguarda esclusivamente i prestiti, in quanto percepiti appunto come non nativi. Al di fuori delle parole che iniziano con (*s*)*wV*, tuttavia, lo schema è identico a quello previsto e osservato per il lessico nativo. Abbiamo infatti *il timer*, *il computer*, *il flyer*, *il trend*, ossia *il* + attacco sillabico, *l'upgrade*, *l'audience*, *l'outlet*, *l'intelligence*, corrispondenti a *l'* + nucleo sillabico, e infine *lo spread*, *lo sponsor*, *lo stress*, *lo show*, *lo shopping*, corrispondente a *s* più consonante o a geminata intrinseca.

⁹ Alla rarità si oppone un'altissima frequenza d'uso: si tratta di parole del lessico di base. Possiamo poi aggiungere la voce *uopo*, che mostra comunque lo stesso schema; cfr. la locuzione *all'uopo*. Segnaliamo infine il sostantivo femminile *uosa*.

¹⁰ Come si vedrà nel paragrafo 6, *swahili* seleziona talvolta anche l'allomorfo *il*.

Ad ogni modo, nel lessico nativo esiste un certo grado di variazione diatopica e probabilmente diafasica e diastratica che ridurrebbe il potere diagnostico della selezione dell'articolo. È noto infatti che nelle varietà settentrionali di italiano di fronte a /sw/¹¹, soprattutto in *suocero*, si seleziona talvolta l'allomorfo *lo*, in particolare da parte di parlanti con un basso livello di istruzione, nonostante la deviazione dallo standard sia stata osservata anche in parlanti/scriventi con un alto livello d'istruzione (Regis, 2017; Leone, 2002). In riferimento alla varietà di italiano parlata in Emilia-Romagna, Leone (2002: 41-42) sostiene che la selezione di *lo* davanti a /sw/ è dovuta alla realizzazione consonantica del *glide*, supponiamo [v] o l'approssimante labiovelare [ʋ] (Gili Fivela, 2010). Sempre in riferimento alle varietà settentrionali, Bertinetto (1999: 90) osserva che talvolta *il* ricorre davanti a *s* impura e riporta l'esempio *il sport* riscontrato nel romanzo di Aldo Busi *Vita standard di un venditore di collant* in bocca a uno dei personaggi di origine settentrionale. Bertinetto (1999: 90, nota 28) sostiene che la selezione di *il* potrebbe essere dovuta a un influsso del dialetto nel quale sono assenti allomorfi terminanti in vocale. Alla luce di ciò, pare che non sia lo schema di selezione dell'articolo a essere soggetto a variazione diatopica e/o diafasico-diastratica, quanto l'interpretazione fonologica e la realizzazione fonetica di suV o il fatto che *lo* non sia incluso nell'inventario allomorfico del parlante o del gruppo di parlanti, estendendo la selezione di *il* anche davanti a quei contesti in cui non ricorrerebbe se il parlante avesse a disposizione un allomorfo che permette una risillabificazione migliore, come *lo*.

Nonostante le considerazioni di cui sopra necessitino di una verifica sistematica, esse suggeriscono che l'osservazione di Canalis (2018) per cui la selezione dell'articolo sembra non essere più una diagnostica valida, data la duplice possibilità osservata di fronte a *wV*, andrebbe riconsiderata: tanto nel lessico nativo quanto nei prestiti, i parlanti sembrano agire con regolarità. La deviazione rispetto al lessico nativo pare riguardare, *ceteris paribus*, solo i dittonghi di tipo *wV*. Nel paragrafo successivo, la questione dell'interpretazione fonologica dei prestiti è trattata in maniera più approfondita.

¹¹ Talvolta anche di fronte a /sj/, come in *lo siero*, nesso che tuttavia sarebbe realizzato [ʃ] e che dunque imporrebbe la selezione di *lo* (REGIS, 2017: 154).

4. *La fonologia dei prestiti in italiano*

In studi dedicati alla fonologia dei prestiti in italiano, per lo più provenienti dall'inglese, è stato osservato che i prestiti, sia antichi che recenti e sia quelli perfettamente integrati che parzialmente integrati, sono adattati secondo la fonologia dell'italiano al momento del loro ingresso nella lingua ospite (Repetti, 1993; Morandini, 2007). Questo è particolarmente valido per i tratti soprasegmentali, come la struttura metrica, vale a dire la sillabificazione e l'assegnazione dell'accento lessicale, che nei prestiti segue generalmente lo schema italiano (Repetti, 1993)¹². Per quanto riguarda la struttura sillabica, soprattutto dei prestiti recenti, Repetti (1993) osserva che, se sul piano ortografico è presente una consonante geminata, questa è presente anche nella realizzazione acustica, quindi fonologica, come in ['flip:er] *flipper* opposto a [z'lo:gan] *slogan*. Si sostiene pertanto che la struttura della sillaba, chiusa (CVC) o aperta (CV, in sillaba atona, e CVV, in sillaba tonica), è dipendente dalla rappresentazione ortografica¹³. Sotto il profilo segmentale, Morandini (2007) osserva che i fonemi nei prestiti che sono presenti nell'inventario italiano vengono trasferiti come tali nella rappresentazione soggiacente e realizzati secondo le regole italiane di derivazione della forma fonetica¹⁴, mentre i fonemi assenti nella fonologia italiana vengono eliminati o sostituiti. Nel processo di sostituzione la tendenza è di associare a un grafema nel prestito il suono che vi corrisponderebbe secondo la conversione grafo-fonemica italiana. Ad esempio, <th>, che in inglese corrisponde alle fricative dentali sorda e sonora /θ ð/, in italiano

¹² Esistono tuttavia delle eccezioni. In *Amsterdam*, ad esempio, la pronuncia italiana può prevedere anche l'accento protosillabico, cioè *Amsterdam*. In realtà, questa struttura accentuale è possibile anche in nederlandese e inglese. Un altro esempio è *continental*, con spostamento dell'accento dalla penultima (in inglese) alla terzultima (in italiano), forse per influsso del presunto accento protosillabico dell'inglese generalizzato dai parlanti italiani (cfr. MAROTTA, 1999: 98).

¹³ Per conclusioni analoghe si vedano i lavori di VANDELIN e PEPERKAMP (2006); MORANDINI (2007); BASSETTI (2017); HAMANN e COLOMBO (2017).

¹⁴ Ad esempio, le occlusive sorde /p t k/ in sillaba tonica non sono aspirate, la consonante rotica /r/ è realizzata come poli- o monovibrante, la laterale /l/ in coda di sillaba non è velarizzata, ecc.

è trasferito come occlusiva sorda /t/ e realizzato [t] a prescindere dal valore di sonorità della fricativa nella lingua d'origine¹⁵, come in [ˈtril:er] *thriller*, anche a causa del valore fonologicamente e foneticamente nullo di <h>, mentre gli italofoeni L1 apprendenti inglese L2 in contesto orale tendono, almeno nelle prime fasi di apprendimento, a sostituire la fricativa dentale sorda con [f] e quella sonora con [v], ossia con i suoni acusticamente più vicini presenti nell'inventario segmentale italiano (Morandini, 2007). È stato proposto che i prestiti sono soggetti alla cosiddetta 'reading adaptation' (Vandelin e Peperkamp, 2006). Con questo termine si fa riferimento al frequente adattamento fonologico sulla base della forma scritta nella lingua da cui il prestito proviene, che viene pronunciata secondo le regole di conversione della lingua di arrivo. In casi del genere, è la fonologia che si adegua allo scritto e non viceversa. Al di fuori dell'italiano, un caso di *reading adaptation* occorrerebbe in portoghese europeo. Le affricate pre-palatali sorda e sonora /tʃ dʒ/ in prestiti dall'inglese sono realizzate come fricative pre-palatali [ʃ ʒ] in, ad es., *check-in* e *jogging*, mentre in parole come *match point*, *Swatch* e *ketchup*, al contrario, l'affricata inglese /tʃ/ è trasferita come tale in portoghese europeo (Vigário, 2021: 153-154). È stato proposto che l'alternanza osservata possa dipendere dalla forma scritta dei prestiti. Limitando la discussione all'affricata sorda /tʃ/, Vigário (2021: 154) ipotizza che la resa fonetica in fricativa [ʃ] emerga a causa della forma scritta <ch>, digramma sempre associato a /ʃ/ in portoghese europeo, mentre l'affricata sorda [tʃ] emergerebbe quando la forma ortografica del prestito prevede il trigramma <tch> nella rappresentazione di /tʃ/: la presenza del grafema <t>, che indica un'occlusiva, implicherebbe la realizzazione della fase di chiusura dell'affricata (Vigário, 2021: 154). Non vogliamo certo sostenere che tutti gli adattamenti facciano riferimento alla rappresentazione ortografica del prestito. Infatti, ad esempio, un fenomeno tipico della fonologia dei prestiti in italiano che prescinde dalla forma ortografica è l'inserimento paragogico

¹⁵ Ad esempio, l'elemento *rutherfordio*, dal nome del chimico e fisico Ernest Rutherford, che nella forma inglese prevede la fricativa dentale sonora /ð/ : <th>, è realizzato in italiano con /t/ : <th>.

di [ə] o inserzioni vocaliche simili in posizione finale se la parola termina con consonante, soprattutto se C finale è una ostruente (*gap* > [ˈgap:ə], *fard* > [ˈfardə]; Broniś, 2020). L'epitesi in *schwa* o simile *intrusive vowel* potrebbe essere dunque una strategia riparatoria di una struttura illecita in italiano, ossia una consonante in posizione finale (Broniś, 2020: 173ss.)¹⁶.

Le osservazioni appena discusse sembrano suggerire che la fonologia dei prestiti sia almeno in parte analoga a quella della lingua nativa; in particolare, la loro rappresentazione fonologica potrebbe essere influenzata dalla forma ortografica, almeno quando, *ça va sans dire*, i prestiti siano entrati a far parte della competenza lessicale dei parlanti, nel nostro caso, italofoeni crucialmente tramite la lingua scritta. Ciò non implica che l'adattamento fonologico dei prestiti segua sempre e sistematicamente i modelli nativi. Un esempio in questo senso è costituito dall'adattamento nel dialetto sardo campidanese dell'affricata pre-palatale sorda /tʃ/ in posizione iniziale proveniente da prestiti catalani, spagnoli e italiani (Lai, 2021: 95-99). Se nel lessico nativo il suono si comporta come le altre ostruenti sorde, vale a dire che subisce il passaggio a fricativa sonora omorganica quando si trova tra due vocali, o acquisisce solo il tratto di sonorità, a seconda della varietà campidanese considerata (V # /tʃ/V → # [ʒ]/[dʒ]V, come in [tʃ] *inisu* "cenere" → *su* [ʒ]/[dʒ] *inisu* "la cenere"), nei prestiti lo stesso suono non subisce processi di lenizione, dal momento che emerge come [tʃ] e viene trattato dai parlanti campidanesi come una consonante soggiacentemente geminata¹⁷, ad esempio in [tʃ] *inesu* "cinese" → *su* [tʃ] *inesu* "il cinese" (Lai, 2021: 95).

Tornando al tema di questo studio, in merito alla selezione dell'articolo davanti ai prestiti con *wV* o *swV* iniziali, l'ipotesi proposta nel paragrafo 1 riguardo l'influenza ortografica sulla classificazione di /w/ come consonante o vocale sarebbe corroborata dalle

¹⁶ Non consideriamo in questa sede la geminazione della consonante finale, perché non pertinente rispetto agli obiettivi di questo studio; per dettagli si rinvia a BAFILLE (2005), PASSINO (2008), e BRONIŚ (2020).

¹⁷ Per la diagnostica utilizzata al fine di determinare lo status di consonante geminata dell'affricata pre-palatale sorda /tʃ/ da prestiti catalani, spagnoli e italiani in sardo campidanese si veda LAI (2021: 97-98).

analisi di Repetti (1993) e Morandini (2007), tra gli altri. In altre parole, il valore consonantico del grafema <w> che rappresenta il *glide* /w/ si trasferirebbe nella rappresentazione fonologica del prestito che lo include, favorendo l'interpretazione consonantica di /w/. Nel lessico nativo lo stesso suono ha valore vocalico (Marotta, 1987; 1988; Canalis, 2018), secondo Regis (2017) a causa del grafema <u> utilizzato nell'ortografia italiana. Sotto un altro punto di vista, si potrebbe ritenere che anche nel lessico nativo il *glide* labiovelare /w/ ha valore consonantico e che la selezione di *l'* davanti agli unici tre lessemi che iniziano con un dittongo *wV* (*uomo* e *uovo*, ed il più arcaico *uopo*) sia dovuta a processi di lessicalizzazione. Sullo statuto vocalico di /w/ si è già discusso nel paragrafo 2, sulla base degli studi di Marotta (1988) e Canalis (2018); si potrebbe aggiungere che la selezione di *l'*, che permette una risillabificazione post-lessicale ottimale, in quanto produce la sillaba [lwɔ]mo/vo, come in [lwɔ]go, non è ristretta all'articolo determinativo. Lo stesso schema si osserva per le preposizioni articolate, *nell'uomo* e *dell'uomo*, per gli aggettivi come *quell'uomo*, *quest'uomo* e *bell'uomo*, forme in linea con l'interpretazione vocalica di /w/, date le quali l'ipotesi della lessicalizzazione risulta meno probabile. Inoltre, la laterale in [lwɔ] è geminata in contesto di Radoppiamento Fonosintattico, come in *sopra*[l'lwɔ]go, fatto che indica che la laterale /l/ davanti al dittongo è in posizione di attacco e non di coda sillabica.

L'interpretazione consonantica di /w/ nei prestiti, opposta a quella vocalica nel lessico nativo, è ulteriormente sostenuta dalla sonorizzazione di /s/ che precede il legamento. Nel lessico nativo la stringa /sw/ è realizzata [sw], con la sibilante sorda ([s]uono, [s]uolo), in linea con l'interpretazione vocalica, quindi nucleare, del legamento e della sibilante come segmento in posizione di attacco. Nei prestiti in cui la stringa è scritta <sw> la sibilante è foneticamente sonora [z], come in ['zwiŋg] *swing* e ['zwɔtʃ] *Swatch*, nonostante nella lingua di origine sia sorda [s] (Janni, 1992). La sonorizzazione di /s/ riguarda tutti i casi in cui questa precede una consonante sonora, come in [z'lb:gan] *slogan*, fenomeno tipico della fonologia italiana. In posizione iniziale, la sonorizzazione di /s/ occorre nei casi di *s* impura, mentre la sibilante prevocalica non

è soggetta al fenomeno. L'esito sonoro [z] di /s/ in *swing*, *Swatch* e nei prestiti analoghi indica pertanto che /w/ è analizzato come consonante in attacco e /s/ come *s* impura. L'interpretazione consonantica di /w/ non sembra essere una caratteristica dei prestiti in quanto tali, ma solo di quelli in cui /w/ è scritto <w>. Nei prestiti in cui il legamento è rappresentato da grafemi vocalici, la sonorizzazione non occorre, come nel prestito dal francese [swa're] *soirée* o dallo spagnolo ['swa:re(t)s] *Suarez*¹⁸. Nel paragrafo successivo si discutono i dittonghi con legamento palatale jV.

5. La selezione dell'articolo davanti a jV: analisi sui corpora

I dittonghi di tipo jV selezionano *lo* a prescindere dalla provenienza del lessema. Tuttavia, in certi prestiti non mancano casi in cui davanti a <y> : /j/ si sia selezionato *il*. Da una ricerca su *Google* emergono numerosi casi del genere, ad esempio *il yogurt naturale funziona molto bene sulla pelle sensibile*, oppure *buon cibo, ma potevano essere più generosi con il yogurt*. Nel lessico nativo, inoltre, è nota l'alternanza tra *lo* e *l'* davanti a /j/: *l'ugoslavo* e *lo iugoslavo* (Vanelli, 1992: 62, nota 70), mentre *l'ieri* sembra essere la forma predominante. Per quanto riguarda quest'ultima parola, si veda ad esempio il grafico nella Figura 4, creato con lo strumento *Ngram Viewer* di *Google*¹⁹, in cui presentiamo i risultati relativi alla parola *ieri* in riferimento a un arco temporale di oltre due secoli, dal 1800 al 2019.

¹⁸ *Suarez* è realizzato da alcuni parlanti con uno iato [su'a:re(t)s].

¹⁹ Lo strumento permette di osservare all'interno del vastissimo corpus di *Google Libri* (il più grande finora esistente a livello globale) la frequenza di determinate stringhe (parole, sintagmi, ecc.) nei testi ivi contenuti: è sufficiente inserire le stringhe che si intende ricercare e lo strumento dà in output un grafico della loro frequenza percentuale nei testi presenti nell'arco di tempo stabilito all'inizio della *query*. Si stimano in totale circa quaranta milioni di testi e cinquecento miliardi di parole, anche se non risulta possibile determinare il numero di testi/parole per lingua (ZANETTIN e PROIETTI, 2021: 168). Lo strumento non fornisce il valore assoluto del dato considerato, ma ne restituisce la frequenza percentuale.

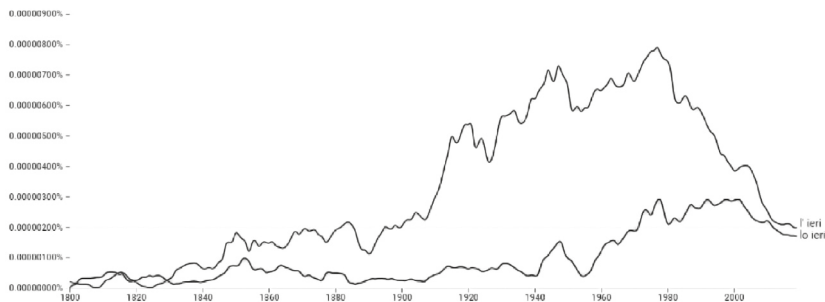


Figura 4. *Frequenza di occorrenza di lo ieri e l'ieri all'interno del corpus Google Books.*

Come si può notare, le due forme cooccorrono, almeno a partire dall'Ottocento. La forma con *l'* supera in frequenza quella con *lo* a partire all'incirca dal 1830, distanziandosi nettamente tra gli anni Dieci e Novanta del ventesimo secolo, per tornare ad avere una frequenza simile dopo il 2010. Ciononostante, *l'ieri* risulta più frequente di *lo ieri* fino al 2019.

Considerando invece *iodio*, la tendenza sembra opposta. Lungo tutto l'Ottocento (lo iodio fu scoperto nel 1811 da Bernard Courtois), davanti a questa parola ricorre quasi esclusivamente l'articolo solo consonantico *l'*. A partire da circa il 1890 *lo iodio* guadagna molto terreno rispetto alla forma concomitante, la quale inizia a essere in disuso. Se tra i termini di ricerca aggiungiamo la grafia ormai desueta *jodio* contestualmente alla forma, prima non considerata, *il iodio*, emerge un quadro interessante. Si vedano le Figure 5 e 6.

Dall'andamento illustrato nelle Figure 5 e 6 possiamo trarre le seguenti conclusioni. Lungo tutto l'Ottocento, la forma più utilizzata era quella con il grafema vocalico iniziale (*iodio*) preceduta dall'allomorfo consonantico *l'*. Altre possibilità coesistevano, ma erano molto meno frequenti. Sorprendentemente, a cavallo tra gli anni Venti del XIX secolo la forma *lo iodio* era assai meno comune di *il iodio*. A partire dall'ultimo decennio del 1800 tutte le forme tranne *l'iodio*, *lo iodio* e *lo jodio* diventano estremamente marginali. Dalla metà degli anni Trenta del XX secolo anche *il iodio* e *lo jodio* perdono drasticamente terreno, cedendo quasi completamente il campo a *lo iodio*. Una situazione analoga si ha per *Ionio*. Se si inserisce questa voce preceduta da ognuno dei tre

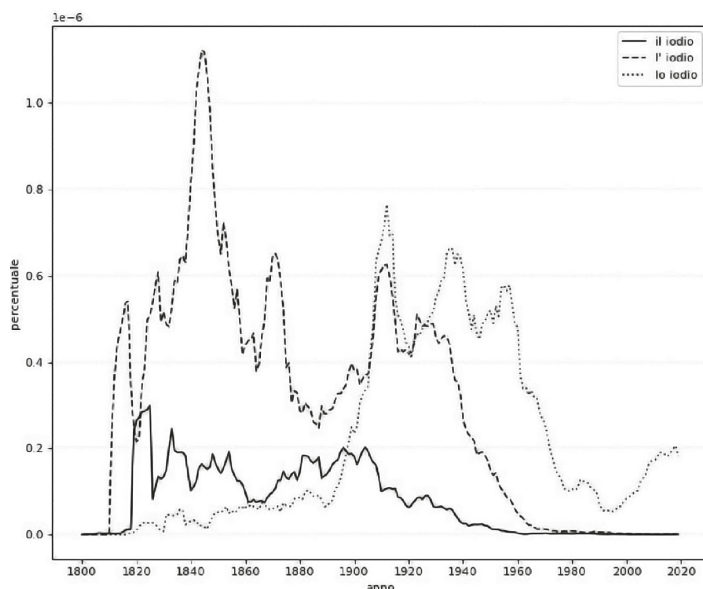


Figura 5. Frequenza di occorrenza di iodio preceduto dai tre allomorfi dell'articolo determinativo maschile singolare.

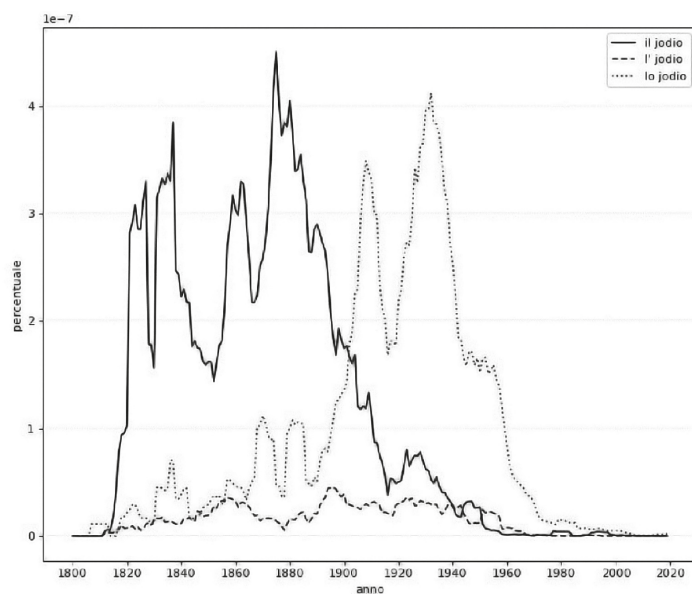


Figura 6. Frequenza di occorrenza di jodio preceduto dai tre allomorfi dell'articolo determinativo maschile singolare.

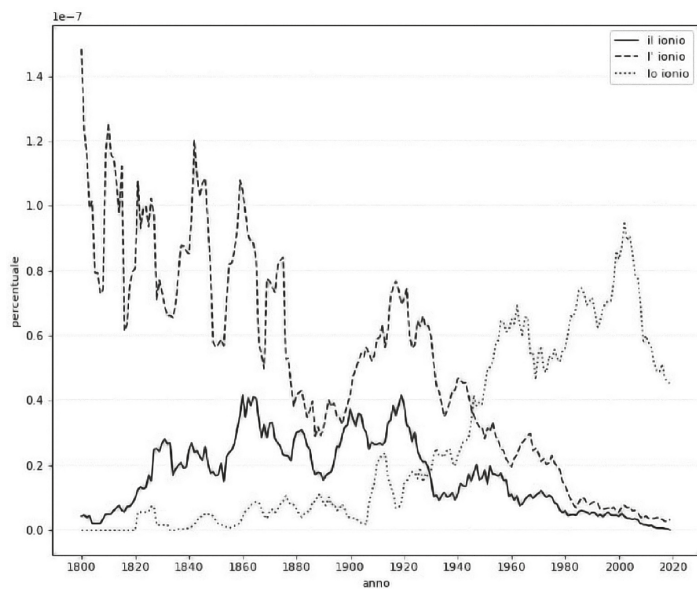


Figura 7. *Frequenza di occorrenza di Ionio preceduto dai tre allomorfi determinativi.*

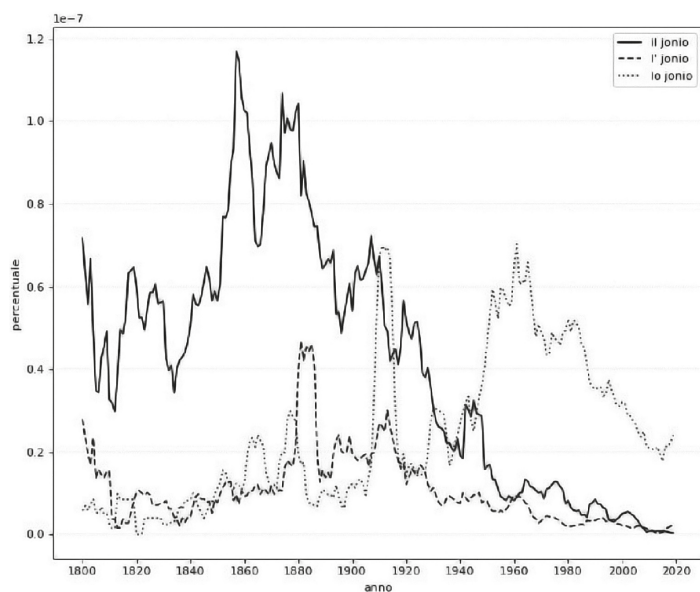


Figura 8. *Frequenza di occorrenza di Jonio preceduto dai tre allomorfi determinativi.*

allomorfi e lo stesso lo si fa per il suo allografo *Jonio*, l'andamento risultante è altrettanto controverso. Si vedano le Figure 7 e 8.

Nelle Figure 7 e 8 vediamo che *l'Ionio* era la forma prevalente fino agli anni Cinquanta del XIX secolo, considerando contestualmente anche la variante con <J>, e fino agli anni Quaranta del XX secolo se consideriamo solamente la forma con <I> iniziale. La seconda forma più comune era *il Jonio* e la terza *il Ionio*. Considerando solo l'allografo con <J> iniziale, *il Jonio* rimane la forma più frequente fino ai primi decenni del Novecento. Le altre forme, e crucialmente quelle con l'articolo vocalico *lo*, sono molto rare lungo tutto l'Ottocento. Una svolta importante si osserva agli inizi del Novecento, periodo in cui *lo Ionio* e *lo Jonio* aumentano la propria frequenza nei testi pubblicati, mentre dagli anni Cinquanta del Novecento diventano predominanti e le altre forme cadono in un progressivo ma rapido disuso.

Alla luce di questi dati, che concernono la selezione dell'articolo davanti a dittonghi *jV*, l'anomalia rappresentata dalle parole che iniziano con il legamento palatale [j], selezionanti *lo*, pare essere un fatto relativamente recente. Inoltre, limitatamente ai testi dell'Ottocento, considerando separatamente le coppie *iodio*, *Ionio* e *jodio*, *Jonio* risulta che per la prima l'allomorfo prevalente era *l'*, mentre per la seconda era *il*. Nonostante la variazione osservata, la distribuzione degli allomorfi davanti a *jV* sembra correlarsi alla grafia di [j], dato che è difficile supporre realizzazioni fonetiche differenti per la stessa parola. Inoltre, <j> era il grafema comunemente usato soprattutto nel XVIII secolo e, anche se in misura minore, nel XIX secolo (Calamai, 2011) per indicare l'approssimante [j]. Sul perché nel corso del Novecento si sia imposta la selezione di *lo* davanti a [j] si può solo speculare. Tuttavia, a meno di ammettere che la struttura sillabica di queste parole sia cambiata drasticamente nell'arco di due secoli, eventualità che non può essere scartata a priori, i dati mostrati suggeriscono che la selezione dell'articolo *lo* davanti a parole con [j] in prima posizione non sia motivata fonologicamente. Si è visto che le forme in assoluto maggioritarie nel XIX secolo erano quelle con *l'* davanti a <iV>. Se i dittonghi *jV* scritti <iV> devono essere analizzati come sequenze nucleari, sia che si parli di nuclei complessi (Marotta, 1987) che di vocali complesse (Canalis, 2018), la scelta dell'allomorfo consonantico *l'* è in linea con lo schema

italiano di selezione dell'articolo. Al contrario, se la resa grafica di /j/ corrispondeva al grafema <j> l'allomorfo prevalente era *il*. Fornire una spiegazione di questa pluralità di *pattern* non sembra possibile da un punto di vista prettamente fonologico.

Dato che l'allomorfo *il* è comunemente selezionato di fronte a attacchi sillabici, semplici o complessi, e visto che <j> è un grafema consonantico utilizzato per rappresentare un suono che non è/era percepito come una vocale vera e propria, la selezione di *il* davanti a <j> è giustificata in quanto in linea con l'analisi del dittongo come una sequenza CV (con C in attacco e V nel nucleo).

Come si vede nelle Figure 5-8, a partire dal Novecento, l'allomorfo *lo* davanti al dittongo *jV* è selezionato in misura molto maggiore rispetto agli altri allomorfi, a prescindere dalla resa ortografica dell'approssimante. Di fatto, se il *glide* [j] è scempio²⁰ (Bertinetto e Loporcaro, 2005), la selezione di *lo* devia dallo schema osservato per gli articoli italiani a prescindere che il *glide* occupi la posizione di attacco sillabico o risieda nel nucleo. È probabile che il fattore che ha governato l'imporsi di *lo* davanti a parole inizianti con /j/ non sia, come si diceva, di carattere fonologico. È possibile infatti che la selezione dell'articolo *lo* davanti a /j/ nell'italiano odierno, comunque non del tutto generalizzata, come rivelato dalla maggior frequenza di *l'* rispetto a *lo* davanti a *ieri* (si veda Figura 4), sia dovuta a una norma prescrittiva fonologicamente immotivata²¹. Fornara (2018), in una rassegna delle grammatiche ottocentesche, a proposito dell'uso degli articoli che queste prescrivevano, segnala che i primi a indicare *lo* davanti alla semiconsonante /j/ furono Morandi e Cappuccini (1894: 75), mentre precedentemente lo stesso allomorfo era indicato solo davanti a *s* impura, <gn>, <gl> e <z>²². Considerato che *lo* davanti a /j/ prende il sopravvento a partire dagli inizi del Novecento, diviene plausibile ammettere che ciò sia dovuto al successo di determinate grammatiche di fine Ottocento, come quella citata, le cui norme sono state mantenute in quelle successive.

²⁰ Tranne nell'italiano parlato a Roma, in cui è lungo (BERTINETTO e LOPORCARO, 2005).

²¹ Oppure basata su varietà diatopiche in cui /j/ è geminato, come quella romana (BERTINETTO e LOPORCARO, 2005).

²² In PUOTI (1839, I: 19), davanti a <z> è accettato anche l'allomorfo *il*.

6. *La selezione dell'articolo davanti a s(w)V: analisi sui corpora*

Se l'ipotesi per cui l'ortografia ha un ruolo attivo nella selezione dell'articolo è corretta, come si evincerebbe dalla distribuzione allomorfica davanti al dittongo *jV* scritto <jV> nell'Ottocento, si potrà spiegare anche la distribuzione degli articoli davanti a (s)*wV* nell'italiano moderno. È stato già suggerito (Janni, 1992; Baroni, 2020) che la duplice grafia con cui è possibile rendere nello scritto l'approssimante labiovelare [w], una propria del lessico nativo, ossia <u>, l'altra proveniente dai prestiti, per lo più inglesi, vale a dire <w>, possa regolare la selezione dell'articolo. A proposito dei nessi /sw/ : <sw>, Janni (1992) infatti scrive:

[il] normale parlante italiano è intimamente convinto che la *w* rappresenti (anzi “sia”) una consonante, come in *Walter*, e che solo per una convenzione “straniera” si debba pronunciarla come la semivocale di *uomo*. Di conseguenza, anche chi non parlerebbe mai di *svòcc* in riferimento al noto orologio, crede in fondo, in una specie di livello subconscio, che *Swatch* non cominci come *suocera*, *suora*, ecc. ma come *svogliato* o *svolazzo*. (Janni, 1992)

Una situazione analoga si avrebbe anche in francese. Se [w] è reso dal digramma vocalico <oi> : /wa/, si seleziona *l'* (*l'oiseau* “l'uccello”), mentre nei prestiti, se reso da <w>, si seleziona *le* (*le wasabi* “il wasabi”). L'ipotesi dell'influenza ortografica sulla scelta dell'allomorfo, applicabile sia all'italiano che al francese, è stata testata da Baroni (2020). Il compito richiesto nello studio, che ha coinvolto parlanti nativi di italiano e di francese, prevedeva di inserire l'articolo determinativo ritenuto più appropriato davanti agli stimoli presentati. Per ogni lingua sono stati elaborati quattro gruppi di parole sulla base della forma scritta dei legamenti. In italiano, i primi due gruppi erano parole e non parole con [w] in prima posizione, rappresentato con il grafema consonantico <w>, gruppo (a), o con quello vocalico <u>, gruppo (b); negli altri due gruppi il suono iniziale era invece [j], rappresentato con <j> o <y>, gruppo (c) o con <i>, gruppo (d). In francese, il gruppo (a) aveva le stesse caratteristiche del corrispondente gruppo italiano (/w/ : <w>), mentre nel gruppo (b) /w/ era

reso con <oi> o <ou> + V. I gruppi (c) e (d) francesi erano analoghi a quelli italiani, a parte l'esclusione del grafema <j> dal gruppo (c), corrispondente a /ʒ/ in francese. Ogni gruppo includeva sia parole reali che non-parole.

Dai risultati non sono emerse differenze tra le parole reali e le non-parole, in nessuna delle due lingue. A veicolare le scelte dei partecipanti è risultata la variabile grafemica. Infatti, nel gruppo (a) sono stati selezionati gli allomorfi *il*, in italiano, e *le*, in francese; nel gruppo (b) è stato selezionato in entrambe le lingue l'allomorfo *l'*. Le differenze tra francese e italiano hanno riguardato i gruppi (c) e (d): in francese, è stato selezionato *le* nel gruppo (c) e *l'* nel gruppo (d), mentre in italiano è stato selezionato *lo* (80% dei casi) in entrambi i gruppi, prescindendo quindi dalla resa grafica di [j].

Per quanto riguarda i gruppi (c) e (d) italiani, è interessante notare che i partecipanti che non hanno selezionato *lo* (20% nel totale dei due sottogruppi) sono stati influenzati dal valore consonantico vs. vocalico del grafema. Infatti, davanti ai grafemi <j> e <y> ha prevalso *il* rispetto a *l'* (16% e 3%, rispettivamente), mentre si ha l'andamento opposto davanti al grafema <i> (*l'* = 12%, *il* = 1%).

Non è tuttavia da escludere un certo grado di artificialità nei risultati del test di Baroni (2020). Un modo per evitare un coinvolgimento attivo dei soggetti è la consultazione di vasti corpora. Anche in questo caso, abbiamo consultato il corpus di *Google Libri* tramite lo strumento *Ngram Viewer*, già utilizzato nel paragrafo 5. La ricerca da noi svolta si è concentrata sulle parole con (s)wV in posizione iniziale; l'arco temporale selezionato è compreso tra il 1900 e il 2019. Gli n-grammi selezionati in questa ricerca corrispondono a digrammi il cui primo elemento è un articolo determinativo e il secondo un sostantivo, in accordo per genere maschile e numero singolare. Ciascuna *query* ha avuto in *input* tre digrammi in cui la variabile è il primo elemento, corrispondente a un diverso allomorfo (*il*, *lo*, *l'*), ad esempio <il uomo, lo uomo, l'uomo>. Ogni *query* è differenziata sulla base del secondo elemento, per un totale di tredici interrogazioni pari al numero dei nomi selezionati in questo studio. Le *query* si dividono in due macro-gruppi: uno in cui [w] è reso dal grafema <u> e un altro in cui [w] è

rappresentato da <w>. A loro volta, questi si distinguono poi per la presenza o assenza della sibilante /s/ in posizione iniziale. Abbiamo quindi quattro gruppi totali:

- (a) il legamento corrisponde a <u> nella grafia e si trova in posizione iniziale; le parole scelte sono obbligatoriamente *uomo* e *uovo*²³;
- (b) il legamento è rappresentato da <w> e occupa la prima posizione; abbiamo scelto per questo contesto *whisk(e)y*, *week(-)end*, *walkman*, *wolof*, *wok*²⁴;
- (c) il legamento, rappresentato da <u>, è preceduto dalla sibilante /s/; i membri scelti per questo gruppo sono *suono*, *suocero*, *suolo*;
- (d) il legamento, rappresentato da <w>, è preceduto dalla sibilante /s/; abbiamo limitato la ricerca a *Swatch*, *swing*, *swahili*.

Davanti ai lessemi nel gruppo (a), l'allomorfo di gran lunga più ricorrente è *l'*. Davanti a queste due parole, *lo* e *il* sono estremamente rari, con una frequenza inferiore allo 0,0000001%.

Passando al gruppo (b), notiamo invece la totale assenza di allomorfi diversi da *il*. Si osservi che anche davanti ai lessemi che iniziano con il dittongo /wɔ/ si presenta lo stesso scenario. Il primo, *walkman*, si riferisce al lettore di musicassette prodotto a partire dal 1978 dalla *Sony*, anno in cui inizia infatti a comparire tra i testi pubblicati. Il secondo lemma, *wok*, è un prestito recente che entra in Italia apparentemente dalla fine degli anni Settanta. Di particolare interesse è l'etnonimo e glottonimo *wolof* in quanto, a differenza dei precedenti, presenta il dittongo /wɔ/ in sillaba tonica aperta. Come nota finale, la forma americana *whisky* risulta più frequente di quella britannica *whiskey*, soprattutto dalla seconda metà del secolo scorso e in particolare dagli anni 2000. A partire dallo stesso periodo, la grafia *weekend* risulta più frequente di *week-end*. I risultati relativi alle voci con /wɔ/ iniziale sono presentati nella Figura 9.

²³ A parte *uopo*, non ci sono altre parole con queste caratteristiche nel lessico italiano.

²⁴ Prestito dal cinese che indica una padella dal fondo emisferico piuttosto alto rispetto a quelle adoperate tradizionalmente nella cucina italiana, attualmente molto comune anche nei paesi occidentali.

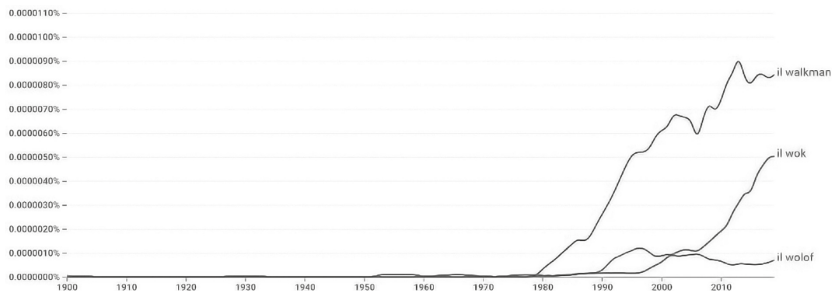


Figura 9. *Frequenza di occorrenza dei lemmi walkman, wok, wolof, ognuno preceduto dagli allomorfi il, lo, l'.*

I risultati relativi al gruppo (b) sono in linea con quelli di Baroni (2020) e possiamo quindi concludere che gli italofoeni si mostrano coerenti nella scelta dell'allomorfo di fronte a parole con /w/ : <w>, selezionando sempre *il*. Davanti a questo gruppo di parole, infatti, non risultano occorrenze con *l'* e *lo*. Al contrario, se il legamento è reso dal grafema vocalico <u> l'allomorfo è *l'*, quello previsto davanti a nuclei.

Passiamo ora al sottogruppo (c), partendo dai lemmi *suono*²⁵ e *suolo*. Il risultato si dà in Figura 10.

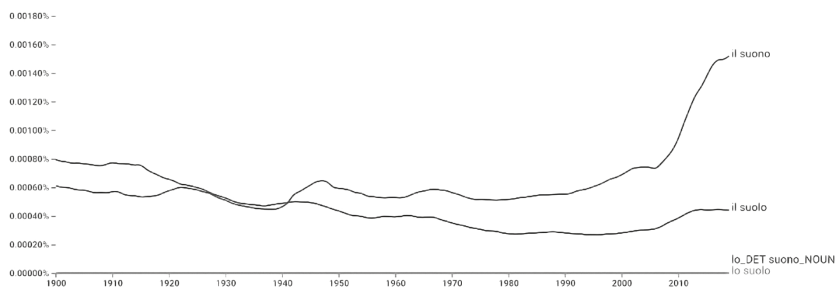


Figura 10. *Frequenza di occorrenza di suono e suolo preceduti dai tre allomorfi il, lo, l'.*

²⁵ Onde evitare confusioni con il pronome omofono *lo* e la prima persona dell'indicativo di *suonare*, abbiamo impostato la ricerca specificando *lo* come determinante e *suono* come nome (lo_DET suono_NOUN), possibilità permessa da *Ngram Viewer*.

Come atteso, *lo suono* è rarissimo²⁶. Inoltre, da una breve consultazione dei testi in cui appare *lo suono*, emerge, ma andrebbe verificato su tutte le occorrenze, che il sintagma nominale in questione proviene prevalentemente da citazioni di versi danteschi, quindi risalenti a una fase diacronica in cui *lo* era usato anche nei contesti in cui ora ricorre soltanto *il* (Vanelli, 1992). Lo scenario mostrato da *suolo* nella Figura 10 è analogo a quello di *suono*. La forma di gran lunga prevalente è quella con *il*. La selezione di *lo* davanti a *suolo* ricorre soprattutto in citazioni dantesche o in ristampe di opere pubblicate prima del XVII secolo. Ad esempio, alcune occorrenze sono presenti nelle ristampe di *Lo cunto de li cunti* di Giambattista Basile (1634), uscite a partire dal 2015. È l'anno in cui è apparsa la trasposizione cinematografica diretta da Matteo Garrone, quindi probabilmente a causa di un rinnovato interesse per la raccolta di racconti da parte del pubblico.

Davanti al terzo membro del gruppo (c), *suocero*, si osserva che nella prima metà del Novecento l'allomorfo *lo* ricorre con maggior frequenza rispetto a quando precede *suono* o *suolo*, seppur con valori percentuali molto bassi (la forma *lo suocero* ha il suo picco nel 1912, raggiungendo lo 0,0000067%). Coerentemente, *lo suocero* risulta attestato ad esempio nelle varietà settentrionali di italiano (Leone, 2002; Regis, 2017; Canalis, 2018), fatto già discusso nella parte introduttiva del paragrafo 4.

Volgiamo ora l'attenzione al sottogruppo (d). Per quanto riguarda *Swatch*, termine che può riferirsi ellitticamente agli orologi prodotti dal brand svizzero, se preceduto da un articolo ricorre con il solo allomorfo *lo*. Le prime occorrenze si hanno a partire dal 1984, l'anno successivo alla fondazione dell'azienda. Parallelamente, *swing*, sin dagli anni Venti, ossia dagli esordi del genere musicale che denota, risulta associato solo all'allomorfo *lo*.

L'andamento degli articoli davanti a *swahili*, etnonimo e glottonimo, è invece altalenante. Fino agli anni Settanta, le due forme, *il swahili* e *lo swahili*, si contendono il terreno. A partire dagli anni Ottanta, tuttavia, *lo swahili* prende nettamente le distanze, affermandosi

²⁶ Trova il suo picco nel 1966, raggiungendo una frequenza dello 0,000003% all'interno del corpus.

come forma più comune, a discapito di *il swahili*, quasi in disuso. Una possibile spiegazione dell'alternanza tra *il* e *lo* davanti a *swahili*, non riscontrata davanti agli altri lessemi con /sw/ iniziale, potrebbe fare appello al fatto che fino agli anni Quaranta del Novecento la forma ortografica più comune reca <u> : /w/, ossia <suahili> (fino agli anni Sessanta le due forme avevano frequenza analoga). Davanti a questa forma, *Ngram Viewer* riporta il solo allomorfo *il*. Questo dato è compatibile con l'analisi di /w/ come consonante solo quando scritto <w>; anche nei prestiti, se /w/ è scritto <u>, la sua interpretazione è vocalica. Per quanto riguarda *il* davanti a *swahili*, notiamo che il periodo in cui ricorre maggiormente è quello in cui la forma con <w> inizia a diffondersi, ma la sua frequenza è simile a quella con <u>. La selezione di *il* potrebbe quindi essere un riflesso dell'abitudine a selezionare *il* davanti a questa parola (fino a quel momento scritta con <u>), tendenza comunque quasi del tutto estinta, in favore di *lo*, l'allomorfo fonologicamente motivato se /w/ ha valore consonantico, una volta che la forma <swahili> si è stabilizzata nel lessico ortografico in italiano.

7. *Discussione*

L'analisi svolta sulla distribuzione degli allomorfi dell'articolo maschile singolare su un vasto corpus mette in luce un grado molto basso di indecidibilità da parte dei parlanti italofoeni nei contesti fonologici esaminati in questo studio, vale a dire parole inizianti con approssimanti, eventualmente preceduti da una sibilante /s/ (si vedano i §§ 1 e 2). I risultati più rilevanti possono essere così riassunti: davanti a parole recanti *wV* in posizione iniziale, dove /w/ è scritto <u>, l'allomorfo selezionato è *l'*. Si tratta di un contesto fonologico-ortografico piuttosto raro nel lessico italiano, rappresentato esclusivamente da *uomo*, *uovo* e la locuzione *all'uopo*, la quale mostra comunque lo stesso schema. Alla rarità del contesto suddetto nel lessico italiano si contrappone l'altissima frequenza d'uso dei primi due lemmi, facenti parte dei circa 2000 lemmi del lessico fondamentale (De Mauro e Cattaneo, 1996). Davan-

ti a parole la cui posizione iniziale è occupata da /w/ scritto <w>, dai dati analizzati nel corpus di *Google Libri* risulta che l'allomorfo selezionato è *il*; la selezione di *il* davanti a questo contesto prescinde dal tipo di sillaba in cui il dittongo si trova, aperta come in *wolof*, chiusa come in *whisky*, e dal timbro della vocale. Di fronte a parole che cominciano con *swV*, dove l'approssimante è reso ortograficamente da <u>, l'allomorfo selezionato è *il*, mentre davanti al gruppo *swV*, il cui approssimante è ortograficamente <w>, si osserva la presenza di *lo*, sistematica davanti a *Swatch* e *swing*, nettamente maggioritaria davanti a *swahili*²⁷.

D'accordo con Baroni (2020), riteniamo che lo scenario qui illustrato (cfr. § 6) possa essere interpretato facendo riferimento cruciale all'ortografia. Dal nostro punto di vista, il legamento /w/ può occupare due posizioni sillabiche diverse a seconda della forma che assume nello scritto. Se /w/ è rappresentato dal grafema consonantico <w>, sarà sillabificato in posizione di attacco, mentre se è scritto con il grafema vocalico <u>, sarà parte di un nucleo complesso o monoposizionale a seconda dell'interpretazione fonologica dei dittonghi di questo tipo.

Un'obiezione all'ipotesi ortografica potrebbe fare appello al fatto che la diversa sillabificazione possa essere motivata fonologicamente senza fare riferimento alla forma scritta. Infatti, i dittonghi in posizione iniziale in alcuni dei prestiti scelti in questo studio (cfr. § 6) sono fonologicamente diversi da quelli di tipo *wV* del lessico nativo, in cui, come abbiamo detto, /w/ può essere seguito da vocali diverse da /ɔ/ solo se preceduto da un'occlusiva velare /k g/. Si potrebbe quindi sostenere che il legamento /w/ in posizione iniziale è interpretato come un membro del nucleo solo se è seguito dalla vocale posteriore /ɔ/, mentre è interpretato come un membro dell'attacco se seguito da altre vocali, giustificando la selezione di *il* davanti a *whisky* e *weekend*. Tuttavia, nei tre prestiti considerati in questo studio che presentano /wɔ/ in posizione iniziale (vale a dire *wok*, *walkman* e *wolof*) è stato comunque selezionato l'allomorfo *il*, mostrando che la qualità della vocale seguente non è pertinente. Inoltre, il dittongo /wɔ/ in *wolof* è perfettamente sovrapponibile a quello in *uomo* e *uovo*, in quanto si trova in sillaba

²⁷ Si veda il paragrafo 6 per una motivazione dell'alternanza tra *il* e *lo* davanti a *swahili*.

tonica aperta, a differenza dei casi rappresentati da *wok* e *walkman* in cui il dittongo è in sillaba chiusa. Nonostante ciò, anche davanti a *wolof* l'allomorfo selezionato è *il*, analogamente agli altri membri del gruppo in questione. A nostro avviso, quindi, il caso di *wolof* corrobora ulteriormente l'ipotesi dell'influsso ortografico sulla selezione dell'articolo davanti al legamento labiovelare.

Dal nostro punto di vista, l'influenza dell'ortografia condiziona la selezione degli allomorfi per il legamento /w/: nel caso in cui /w/ riceva interpretazione consonantica, occuperà la posizione di attacco sillabico; viceversa, se a /w/ è assegnato valore vocalico, il legamento occuperà una posizione nel nucleo sillabico. L'effetto di queste due sillabificazioni si riflette sull'allomorfo selezionato davanti a /w/.

Analogamente, a causa dell'interpretazione consonantica di /w/ in *swing*, *Swatch*, ecc., la sibilante è interpretata in queste parole come un caso di *s impurum*, e dunque l'allomorfo selezionato è coerentemente *lo*, come accade nel lessico italiano; si vedano ad esempio *lo stato*, *lo sparo*, *lo sconto*. La struttura sillabica di parole con (s)wV iniziale, dove /w/ è scritto <w> o <u> è rappresentata nelle Figure 11 e 12.

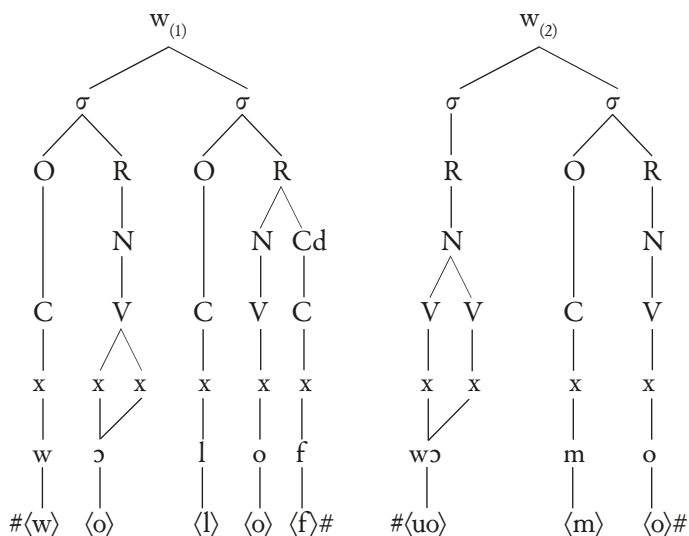


Figura 11. Rappresentazione sillabica delle parole (1) *wolof*, (2) *uomo*.
I dittonghi nucleari sono interpretati qui come monoposizionali.

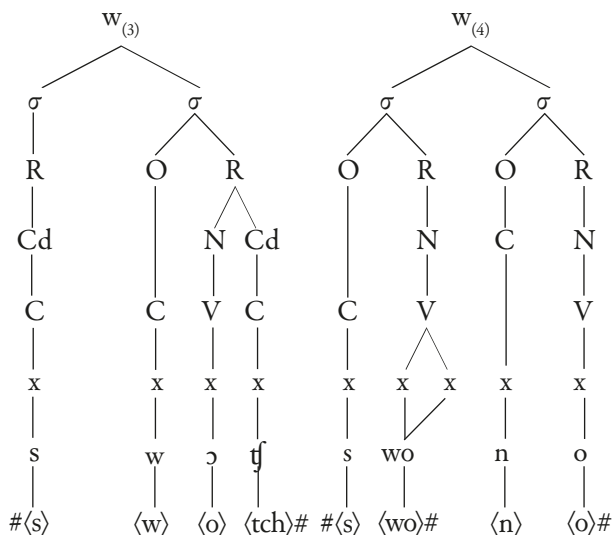


Figura 12. Rappresentazione sillabica delle parole (3) Swatch, (4) suono.
I dittonghi nucleari sono interpretati qui come monoposizionali.

Nel nostro studio, la rappresentazione ortografica sembra avere un ruolo attivo nella rappresentazione fonologica relativamente al fenomeno preso in esame. I nostri risultati sono in linea con quelli discussi da Taft (2006) concernenti il trattamento di /r/ preconsontico nella varietà non rotica di inglese australiano. L'autore ritiene che la rotica sia fonologicamente reale solo a causa della sua presenza ortografica, non essendo foneticamente realizzato in questa posizione, come in [kɔ:n] *corn* "mais" e [fɑ:m] *farm* "fattoria". Un altro studio molto noto che fa appello al ruolo dell'ortografia nella rappresentazione fonologica è quello pionieristico di Ehri e Wilce (1980), in cui i parlanti si mostravano sicuri di 'sentire' un suono aggiuntivo, nello specifico /t/, in parole come *pitch* "tono" e non in lessemi come *rich* "ricco", come si evince dai risultati del *phoneme counting task* da loro eseguito con parlanti anglofoni L1. Evidenze analoghe sono discusse da Seidenberg e Tanenhaus (1979), nel cui studio i parlanti, inglese L1, riscontravano difficoltà nel determinare se la coppia di parole *rye - tie* fosse in rima rispetto alla coppia *pie - tie*, in cui l'individuazione della rima era immediata. Secondo l'analisi

degli autori, la differenza risiederebbe nella forma ortografica delle parole: i parlanti effettuerebbero il confronto su di questa piuttosto che sulla forma fonologica. Bassetti (2017) ha mostrato che i parlanti italiano L1 tendono a produrre in inglese L2 delle geminate in luogo di una consonante ortograficamente doppia ([*'kit:i*] *kitty* vs. [*'si:ti*] *city*), anche in contesto di apprendimento caratterizzato da input orale e scritto. L'influenza ortografica è emersa inoltre in studi che hanno utilizzato compiti di riconoscimento lessicale – ossia determinare se una parola udita è reale o inventata – come Ziegler e Ferrand (1998). Altri studi che hanno osservato un certo grado di influenza ortografica sull'elaborazione di un segnale orale sono Taft e Hambly (1985), Halle *et al.* (2000), Damian e Bower (2003), Ventura *et al.* (2004), tra gli altri.

Studi come quelli appena citati suggeriscono che vada posta maggiore attenzione al ruolo che l'ortografia ricopre a livello fonologico e psicolinguistico. Potrebbe essere pertanto giunto il tempo di riconsiderare, almeno parzialmente, la concezione per cui «writing is not language but merely a way of recording language by means of visible marks» (Bloomfield, 1946 [1933¹]: 21). Nelle società moderne, altamente alfabetizzate, la lingua scritta ha infatti acquisito un ruolo più importante della semplice registrazione di un segnale orale, diventando una manifestazione della facoltà del linguaggio parallela a quella orale: l'una e l'altra rappresentazione si influenzano reciprocamente, interconnettendosi (Rastle *et al.*, 2011: 1592), in modi e intensità ancora da valutare.

Ringraziamenti

Desidero ringraziare i due revisori anonimi di SSL per i loro commenti critici, che hanno contribuito a chiarire le prospettive avanzate in questo articolo. Sono inoltre grato alla prof.ssa Giovanna Marotta per le preziose osservazioni e gli spunti di riflessione.

Bibliografia

- BAFILE, L. (2005), *Le consonanti finali nel fiorentino e nel napoletano*, in «Quaderni del dipartimento di linguistica. Università degli studi di Firenze», 15, 1, pp. 1-25.
- BARONI, A. (2020), *Le semivocali in francese e italiano. Alternanza con le vocali alte e selezione dell'articolo*, in BACQUIN, M., BERNARDINI, P., EGERLAND, V. e GRANFELDT, J. (2020, eds.), *Écrits sur les langues romanes à la mémoire d'Alf Lombard*, Lund University, Lund, pp. 17-42.
- BASSETTI, B. (2017), *Orthography affects second language speech: Double letters and geminate production in English*, in «Journal of Experimental Psychology: Learning, Memory, and Cognition», 43, 11, pp. 1835-1842.
- BERTINETTO, P.M. (1999), *La sillabazione dei nessi /sC/ in italiano: un'eccezione alla tendenza universale?*, in BENINCÀ, P., VANELLI, L. e MIONI, A. (1999, a cura di), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del 31 congresso della Società di linguistica italiana (Padova, 25-27 settembre 1997)*, Bulzoni, Roma, pp. 71-96.
- BERTINETTO, P.M. e LOPORCARO, M. (2005), *The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome*, in «Journal of the International Phonetic Association», 35, 2, pp. 131-151.
- BLOOMFIELD, L. (1946 [1933¹]), *Language*, Holt, Rinehart & Winston, New York.
- BRADLEY, T. (2014), *Optimality theory and Spanish phonology*, in «Language and Linguistics Compass», 8, 2, pp. 65-88.
- BRONIŚ, O. (2020), *Word-final Vowel Epenthesis in Italian Loanword Adaptation*, Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego, Warszawa.
- CALAMAI, S. (2011), *Semivocali*, in *Enciclopedia dell'Italiano Treccani* [accessibile online: [https://www.treccani.it/enciclopedia/semivocali_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/semivocali_(Enciclopedia-dell'Italiano)/), ultimo accesso 11.10.2021].

- CALAMAI, S. e BERTINETTO, P.M. (2006), *Per uno studio articolatorio dei legamenti palatale, labio-velare e labio-palatale dell'italiano*, in GIORDANI, V., BRUSEGHINI, V. e COSI, P. (2006, a cura di), *Scienze vocali e del linguaggio - Metodologie di valutazione e risorse linguistiche. Atti del III convegno nazionale AISV*, EDK Editore, Torriana, pp. 43-56.
- CAMILLI, A. e FIORELLI, P. (1965), *Pronuncia e grafia dell'italiano*, Sansoni, Firenze.
- CANALIS, S. (2018), *The status of Italian glides in the syllable*, in PETROSINO, R., CERRONE, P. e VAN DER HULST, H. (2018, eds.), *From Sounds to Structures: Beyond the Veil of Maya*, De Gruyter Mouton, Berlin / Boston, pp. 3-29.
- CORTELAZZO, M. e ZOLLI, P. (1980), *Dizionario etimologico della lingua italiana*. Vol. 2: *D-H*, Zanichelli, Bologna.
- DAMIAN, M.F. e BOWERS, J.S. (2003), *Assessing the role of orthography in speech perception and production: Evidence from picture-word interference tasks*, in «European Journal of Cognitive Psychology», 21, pp. 581-598.
- DAVIS, S. (1990), *Italian onset structure and the distribution of il and lo*, in «Linguistics», 28, pp. 43-55.
- DE LA CRUZ CABANILLAS, I., TEJEDOR MARTÍNEZ, T., DÍEZ PRADOS, M. e CERDÁ REDONDO, E. (2007), *English loanwords in Spanish computer language*, in «English for Specific Purposes», 26, 1, pp. 52-78.
- DE MAURO, T. e CATTANEO, A. (1996), *DIB. Dizionario di base della lingua italiana*, Paravia, Torino.
- EHRI, L. e WILCE, L. (1980), *The influence of orthography on readers' conceptualization of the phonemic structure of words*, in «Applied Psycholinguistics», 1, pp. 371-385.
- FORNARA, S. (2018), *Pronome e articolo*, in ANTONELLI, G., MOTOLOSE, M. e TOMASIN, L. (2018, a cura di), *Storia dell'italiano scritto IV. Grammatiche*, Carocci, Roma, pp. 261-292.
- GILI FIVELA, B. (2010), *Labiali*, in *Enciclopedia dell'italiano Treccani* [accessibile online: [https://www.treccani.it/enciclopedia/labiali_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/labiali_(Enciclopedia-dell'Italiano)/), ultimo accesso 11.10.2021].
- GOLDSMITH, J.A. (1976), *Autosegmental Phonology*, PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA.

- HALLE, P.A., CHEREAU, C. e SEGUI, J. (2000), *Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds*, in «Journal of Memory & Language», 43, pp. 618-639.
- HAMANN, S. e COLOMBO, I. (2017), *A formal account of the interaction of orthography and perception: English intervocalic consonants borrowed into Italian*, in «Natural Language & Linguistic Theory», 35, 3, pp. 683-714.
- HARRIS, J. e LINDSEY, G. (1995), *The elements of phonological representation*, in DURAND, J. e KATAMBA, F. (1995, eds.), *Frontiers of Phonology. Atoms, Structures and Derivations*, Longman, Harlow, pp. 34-79.
- HUALDE, J.I. (1989), *Silabeo y estructura morfémica en español*, in «Hispania», 22, 4, pp. 821-831.
- JACOBS, H. e VAN GERWEN, R. (2006), *Glide strengthening in French and Spanish and the formal representation of affricates*, in GESS, R.S. e ARTEAGA, D. (2006, eds.), *Historical Romance Linguistics. Retrospectives and Linguistics*, John Benjamin, Amsterdam / Philadelphia, pp. 77-96.
- JANNI, P. (1992), *Lo Swatch*, in «Lingua Nostra», pp. 86-87.
- KAYE, J. (1992), *Do you believe in magic? The story of s+C sequences*, in «SOAS working papers in Linguistics and Phonetics», 2, pp. 293-313.
- KAYE, J., LOWENSTAMM, J. e VERGNAUD, J.R. (1990), *Constituent structure and government in phonology*, in «Phonology», 7, 1, pp. 193-231.
- LAI, R. (2021), *Word-initial heterosyllabic clusters in loanword adaptation: From Catalan, Spanish and Italian loanwords into Sardinian*, in LLORET, M.-R. e PONS-MOLL, C. (2021, eds.), *L'adaptació de manlleus en català i en altres llengües romàniques*, Edicions de la Universitat de Barcelona, Barcellona, pp. 77-110.
- LEONE, L. (2002), *Conversazioni sulla lingua italiana*, Olschki Editore, Firenze.
- LOPORCARO, M. (2011), *Phonological processes*, in MAIDEN, M., SMITH, J.C. e LEDGEWAY, A. (2011, eds.), *The Cambridge History of the Romance Languages*. Vol. 1: *Structures*, Cambridge University Press, Cambridge, pp. 109-154.
- MAROTTA, G. (1985), *Modelli e misure ritmiche: la durata vocalica in italiano*, Zanichelli, Bologna.

- MAROTTA, G. (1987), *Dittongo e iato in italiano: una difficile discriminazione*, in «Annali della Scuola Normale Superiore di Pisa», 17, pp. 847-887.
- MAROTTA, G. (1988), *The Italian diphthongs and the autosegmental framework*, in BERTINETTO, P.M. e LOPORCARO, M. (1988, eds.), *Certamen Phonologicum*, Rosenberg & Sellier, Torino, pp. 389-420.
- MAROTTA, G. (1993), *Selezione dell'articolo e sillaba in italiano: un'interazione totale?*, in «Studi di grammatica italiana», 15, pp. 255-296.
- MAROTTA, G. (1999), *Degenerate feet nella fonologia metrica dell'italiano*, in BENINCÀ, P., MIONI, A. e VANELLI, L. (1999, a cura di), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI congresso della Società di linguistica italiana*, Bulzoni, Roma, pp. 97-116.
- MAROTTA, G. (2010), *Iato*, in *Enciclopedia dell'Italiano Treccani* [accessibile online: https://www.treccani.it/enciclopedia/iato_%28Enciclopedia-dell%27Italiano%29/, ultimo accesso 11.10.2022].
- MORALES-FRONT, A. (2018), *The syllable*, in GEESLIN, K.L. (2018, eds.), *The Cambridge Handbook of Spanish Linguistics*, Cambridge University Press, Cambridge, pp. 190-210.
- MORANDI, L. e CAPPUCCINI, G. (1894), *Grammatica italiana (regole ed esercizi) per uso delle scuole ginnasiali tecniche e complementari*, Paravia, Torino / Milano / Roma / Firenze / Napoli.
- MORANDINI, D. (2007), *The phonology of loanwords into Italian*, MA Thesis, University College London, London.
- NESPOR, M. (1993), *Fonologia*, il Mulino, Bologna.
- PASSINO, D. (2008), *Aspects of Consonantal Lengthening in Italian*, Unipress, Padova.
- PETROSINO, R. (2018), *Allomorphy of Italian determiners at the morphology-phonology interface*, in PETROSINO, R., CERRONE, P. e VAN DER HULST, H. (2018, eds.), *From Sounds to Structures: Beyond the Veil of Maya*, De Gruyter Mouton, Berlin / Boston, pp. 295-328.
- PUOTI, B. (1939), *Regole elementari della lingua italiana*, Tipografia delle Scienze, Roma.

- RASTLE, K., MCCORMICK, S.F., BAYLISS, L. e DAVIS, C.J. (2011), *Orthography influences the perception and production of speech*, in «Journal of Experimental Psychology: Learning, Memory, and Cognition», 37, 6, pp. 1588-1594.
- REGIS, R. (2017), *How standard regional Italians set in: The case of standard Piedmontese Italian*, in CERRUTI, M., CROCCO, C. e MARZO, S. (2017, eds.), *Towards a New Standard. Theoretical and Empirical Studies on the Restandardization of Italian*, De Gruyter, Boston / Berlin, pp. 145-175.
- REPETTI, L. (1993), *The integration of foreign loans in the phonology of Italian*, in «Italice», 70, 2, pp. 182-196.
- SALZA, P.L. (1988), *Durations of Italian diphthongs and vowel clusters*, in «Language and Speech», 31, 2, pp. 97-113.
- SALZA, P.L., MAROTTA, G. e RICCA, D. (1987), *Duration and formant frequencies of Italian bivocalic sequences*, in *Proceedings XIth ICPhS: The Eleventh International Congress of Phonetic Sciences August 1-7 1987 Tallinn Estonia U.S.S.R.*, Academy of Sciences of the Estonian S.S.R., Tallin, 3, pp. 113-116.
- SEIDENBERG, M.S. e TANENHAUS, M.K. (1979), *Orthographic effects on rhyme monitoring*, in «Journal of Experimental Psychology: Human Learning and Memory», 5, 6, pp. 546-554.
- SERIANNI, L. (1998), *Lezioni di grammatica storica italiana*, Bulzoni, Roma.
- SMEAD, R. (1998), *English loanwords in Chicano Spanish: Characterization and rationale*, in «Bilingual Review / La Revista Bilingüe», 23, 2, pp. 113-123.
- TAFT, M. (2006), *Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers*, in «Journal of Psycholinguistic Research», 35, 1, pp. 67-78.
- TAFT, M. e HAMBLY, G. (1985), *The influence of orthography on phonological representations in the lexicon*, in «Journal of Memory & Language», 24, pp. 320-335.
- VAN DER VEER, B. (2006), *The Italian 'Mobile Diphthongs'. A Test Case for Experimental Phonetics and Phonological Theory*, LOT, Utrecht.

- VANDELIN, I. e PAPERKAMP, S. (2006), *The influence of orthography on loanword adaptations*, in «Lingua», 116, pp. 996-1007.
- VANELLI, L. (1992), *Da “lo” a “il”: storia dell’articolo definito maschile singolare nell’italiano e nei dialetti settentrionali*, in «Rivista italiana di dialettologia», 16, pp. 29-66.
- VENTURA, P., MORAIS, J., PATTAMADILOK, C. e KOLINSKY, R. (2004), *The locus of the orthographic consistency effect in auditory word recognition*, in «Language & Cognitive Processes», 19, pp. 57-95.
- VIGÁRIO, M. (2021), *Loanword adaptation and European Portuguese segmental phonology*, in LLORET, M.-R. e PONS-MOLL, C. (2021, eds.), *L’adaptació de manlleus en català i en altres llengües romàniques*, Edicions de la Universitat de Barcelona, Barcellona, pp. 145-174.
- WITUCKI, J. (1976), *The consonants of Chamorro*, in WITUCKI, J., WALROD, M.R. e SHAND, J. (1976, eds.), *Papers in Philippine Linguistics*, The Australian National University, Canberra, pp. 1-20.
- ZIEGLER, J. e FERRAND, L. (1998), *Orthography shapes the perception of speech: The consistency effect in auditory word recognition*, in «Psychonomic Bulletin & Review», 5, 4, pp. 683-689.

PIERO COSSU

Dipartimento di Filologia, Letteratura e Linguistica

Università di Pisa

Via Santa Maria 36

56126 Pisa (Italia)

piero.cossu@fileli.unipi.it

NORME PER GLI AUTORI

Le proposte editoriali (articoli, discussioni e recensioni), redatte in italiano, inglese o altra lingua europea di ampia diffusione, vanno inviate tramite il sistema *Open Journal System* (OJS) collegandosi al sito <http://www.studiesaggilinguistici.it> (ove sono indicate le procedure da seguire), utilizzando due formati: un file pdf anonimo e un file word completo di tutti i dati dell'Autore (indirizzo istituzionale e/o privato, numero telefonico ed e-mail).

Nella redazione della proposta editoriale, gli Autori sono invitati ad attenersi scrupolosamente alle norme redazionali della rivista, disponibili sul sito.

Le proposte di articoli e discussioni dovranno essere corredate da un breve riassunto anonimo in lingua inglese, della lunghezza di circa 15 righe o 1.000 battute (spazi inclusi) e da 3 o 4 parole-chiave che individuino dominio e tema dell'articolo.

I contributi saranno sottoposti alla lettura critica di due *referees* anonimi, e quindi all'approvazione del Comitato Editoriale.

Il contributo accettato per la pubblicazione e redatto in forma definitiva andrà inviato tramite OJS nei tempi indicati dal sistema, sia in formato word che pdf, includendo i font speciali dei caratteri utilizzati.

Edizioni ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com - www.edizioniets.com
Finito di stampare nel mese di maggio 2023